

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 September 2003 (18.09.2003)

PCT

(10) International Publication Number
WO 03/076620 A1

- (51) International Patent Classification⁷: C12N 15/11, C12Q 1/68
- (21) International Application Number: PCT/AU03/00293
- (22) International Filing Date: 12 March 2003 (12.03.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/363,852 14 March 2002 (14.03.2002) US
- (71) Applicant: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION [AU/AU]; Limestone Avenue, Campbell, Australian Capital Territory 2601 (AU).
- (72) Inventors: WATERHOUSE, Peter; 5 Banjine Street, O'Connor, Australian Capital Territory 2602 (AU). WESLEY, Susan; 18 Pambula Street, Kaleen, Australian Capital Territory 2617 (AU). HELLIWELL, Chris; 167 Wattle Street, O'Connor, Australian Capital Territory 2602 (AU).
- (74) Agent: BLAKE DAWSON WALDRON PATENT SERVICES; 39th Floor, 101 Collins Street, Melbourne, Victoria 3000 (AU).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— with international search report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



WO 03/076620 A1

(54) Title: METHODS AND MEANS FOR MONITORING AND MODULATING GENE SILENCING

(57) Abstract: Methods and means are provided for monitoring and modulating reduction of gene expression in eukaryotic organisms, using double stranded RNA comprising, in addition to the dsRNA region comprising nucleotide sequences homologous to the target gene, additional dsRNA regions designed to down.

Methods And Means For Monitoring And Modulating Gene Silencing.

Field of the invention.

5

The present invention relates to methods of altering the expression of genes in eukaryotic organisms, such as plants, but also animals such as nematodes, insects and arthropods, mammals including humans, or yeasts, fungi or molds, using dsRNA capable of altering the expression of target genes, or genes
10 encoding such dsRNA. Also provided are non-human eukaryotic organisms comprising such dsRNA or genes encoding it.

In a first aspect, the invention provides methods and means for monitoring the silencing of a target gene in a eukaryotic cell by measuring the degree of
15 silencing of a second gene wherein the silencing of the target gene and of the second gene is obtained through the action of a single initial dsRNA molecule provided to the eukaryotic cell.

In a second aspect, the invention provides methods and means for modulating
20 the degree of silencing of a target gene in a eukaryotic cell, by including into the single initial dsRNA molecule provided to the eukaryotic cell in addition to the target dsRNA inducing the silencing of the target gene, an amount of dsRNA sequences unrelated to the target dsRNA, in a proportion reflecting the desired modulation of the degree of silencing of the target gene.

25

Background art.

Until recently, two predominant methods for the modulation of gene expression in eukaryotic organisms were known, which are referred to in the art as
30 "antisense" downregulation or "sense" downregulation.

Recent work has demonstrated that the silencing efficiency could be greatly improved both on quantitative and qualitative level using chimeric constructs encoding RNA capable of forming a double stranded RNA by basepairing between the antisense and sense RNA nucleotide sequences respectively
5 complementary and homologous to the target sequences. Such double stranded RNA (dsRNA) is also referred to as hairpin RNA (hpRNA) or interfering RNA (RNAi).

Down-regulating the expression of target nucleotide sequence using dsRNA
10 has been described as a convenient analysis method for the elucidation of the function of nucleic acids the role of which is unknown (see e.g. WO99/53050 or WO00/01846). dsRNA mediated gene silencing can also be used to modulate the expression of one or more target genes in an organism to obtain a modified organism with a desired phenotype or trait (WO99/53050).

15 It would be helpful for the above mentioned applications to have a way for easy monitoring the quantitative and qualitative down-regulation of the target nucleic acids. This is particularly true where determination of the down-regulation of the expression of the target gene through the analysis of the phenotype caused by
20 the down regulation is labor-intensive, cost-intensive, time-consuming, requires the use of specialized analysis methods, etc. . When dsRNA mediated gene silencing is used for function identification/validation of unknown genes, where by definition the phenotype to look for is not known, it may be imperative to have an alternative method for monitoring the efficiency of the silencing.

25 Further, it would be convenient to be able to modulate the degree of silencing of a particular target gene or genes in an organism e.g. by generating a population of organisms exhibiting a large spectrum in the individual degrees of dsRNA mediated gene silencing and identifying the organisms with the desired
30 degree of silencing.

Fire *et al.*, 1998 describe specific genetic interference by experimental introduction of double-stranded RNA in *Caenorhabditis elegans*.

WO 99/32619 provides a process of introducing an RNA into a living cell to
5 inhibit gene expression of a target gene in that cell. The process may be practiced *ex vivo* or *in vivo*. The RNA has a region with double-stranded structure. Inhibition is sequence-specific in that the nucleotide sequences of the duplex region of the RNA and or a portion of the target gene are identical.

10 Waterhouse *et al.* 1998 describe that virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and anti-sense RNA. The sense and antisense RNA may be located in one transcript that has self-complementarity.

15 Hamilton *et al.* 1998 describes that a transgene with repeated DNA, *i.e.*, inverted copies of its 5' untranslated region, causes high frequency, post-transcriptional suppression of ACC-oxidase expression in tomato.

WO 98/53083 describes constructs and methods for enhancing the inhibition of
20 a target gene within an organism which involve inserting into the gene silencing vector an inverted repeat sequence of all or part of a polynucleotide region within the vector.

WO 99/53050 provides methods and means for reducing the phenotypic
25 expression of a nucleic acid of interest in eukaryotic cells, particularly in plant cells, by introducing chimeric genes encoding sense and antisense RNA molecules directed towards the target nucleic acid. These molecules are capable of forming a double stranded RNA region by base-pairing between the regions with the sense and antisense nucleotide sequence or by introducing
30 the RNA molecules themselves. Preferably, the RNA molecules comprise simultaneously both sense and antisense nucleotide sequences.

WO 99/49029 relates generally to a method of modifying gene expression and to synthetic genes for modifying endogenous gene expression in a cell, tissue or organ of a transgenic organism, in particular to a transgenic animal or plant.

5 Synthetic genes and genetic constructs, capable of forming a dsRNA which are capable of repressing, delaying or otherwise reducing the expression of an endogenous gene or a target gene in an organism when introduced thereto are also provided.

10 WO 99/61631 relates to methods to alter the expression of a target gene in a plant using sense and antisense RNA fragments of the gene. The sense and antisense RNA fragments are capable of pairing and forming a double-stranded RNA molecule, thereby altering the expression of the gene. The present invention also relates to plants, their progeny and seeds thereof obtained using

15 these methods.

WO 00/01846 provides a method of identifying DNA responsible for conferring a particular phenotype in a cell which method comprises a) constructing a cDNA or genomic library of the DNA of the cell in a suitable vector in an

20 orientation relative to (a) promoter(s) capable of initiating transcription of the cDNA or DNA to double stranded (ds) RNA upon binding of an appropriate transcription factor to the promoter(s); b) introducing the library into one or more of cells comprising the transcription factor, and c) identifying and isolating a particular phenotype of a cell comprising the library and identifying the DNA

25 or cDNA fragment from the library responsible for conferring the phenotype. Using this technique, it is also possible to assign function to a known DNA sequence by a) identifying homologues of the DNA sequence in a cell, b) isolating the relevant DNA homologue(s) or a fragment thereof from the cell, c) cloning the homologue or fragment thereof into an appropriate vector in an

30 orientation relative to a suitable promoter capable of initiating transcription of dsRNA from said DNA homologue or fragment upon binding of an appropriate

transcription factor to the promoter and d) introducing the vector into the cell from step a) comprising the transcription factor.

WO 00/44914 also describes composition and methods for *in vivo* and *in vitro* attenuation of gene expression using double stranded RNA, particularly in zebrafish.

WO 00/49035 discloses a method for silencing the expression of an endogenous gene in a cell, the method involving overexpressing in the cell a nucleic acid molecule of the endogenous gene and an antisense molecule including a nucleic acid molecule complementary to the nucleic acid molecule of the endogenous gene, wherein the overexpression of the nucleic acid molecule of the endogenous gene and the antisense molecule in the cell silences the expression of the endogenous gene.

15

Smith et al., 2000 as well as WO 99/53050 described that intron containing dsRNA further increased the efficiency of silencing. Intron containing hairpin RNA is often also referred to as ihpRNA.

As illustrated by the above mentioned references, dsRNA mediated gene silencing is a phenomenon which occurs in a wide range of eukaryotic organisms, including plants, yeasts or fungi, insects, arthropods and vertebrate animals, including mammals.

WO 93/23551 describes a process for the inhibition of two or more target genes which comprises introducing into the plant a single control gene which has distinct DNA regions homologous to each of the target genes and a promoter operative in plants adapted to transcribe from such distinct regions either antisense or sense RNA that inhibits expression of each of the target genes.

30

WO 99/49029 describes a method for simultaneously targeting the expression of several target genes which are co-expressed in a particular cell, for example by using a dispersed nucleic acid molecule or foreign nucleic acid molecule which comprises nucleotide sequences which are substantially identical to each
5 of said co-expressed target genes.

However, none of the above mentioned prior art references have addressed the problem of monitoring the degree of silencing of a target gene with a phenotype that is difficult or practically impossible to measure. The prior art
10 also remains defective in providing a method for modulating or fine-tuning the degree of silencing of a specific target gene or genes using dsRNA.

These and other problems have been solved as described hereinafter in the different embodiments and claims.

15

SUMMARY OF THE INVENTION.

The invention provides a method for monitoring the reduction of the expression of a target gene in a cell of a eukaryotic organism, such as a plant, animal,
20 yeast, fungus or mold, comprising the steps of

- a) providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region wherein
 - i) the first region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19
25 consecutive nucleotides from a sense nucleotide region of said target gene;
 - ii) the second region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides
30 from said sense nucleotide region of the target gene ;

- iii) the first region and the second region are capable of forming a double stranded RNA region;
 - iv) the third region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene, such as an endogenous gene or a transgene, stably integrated into the genome of the eukaryotic cell and which is different from the target gene;
 - v) the fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of the 19 consecutive nucleotides from the sense nucleotide region of the second gene;
 - vi) the third and fourth region being capable of forming a double stranded RNA region; and
- b) monitoring the reduction of the expression of the target gene by analyzing the reduction in expression of said second gene.

The dsRNA may be transcribed from a chimeric gene comprised within cells of the eukaryotic organism, the chimeric gene comprising:

- c) a promoter region which functions in the eukaryotic cell; operably linked to
- d) a DNA region which when transcribed yields the dsRNA molecule; and
- e) a transcription termination and polyadenylation region that functions in the cells of eukaryotic organism.

25

It is also an object of the invention to provide an RNA molecule as described as well as the use of such an RNA molecule for measuring the reduction of expression of a target gene.

- 30 It is another object of the invention to provide a DNA molecule for measuring the reduction of expression of a target gene in a eukaryotic cell comprising

- a) a promoter region which functions in the eukaryotic cell; operably linked to
- b) a DNA region which when transcribed yields a dsRNA molecule, wherein the dsRNA comprises a first, second, third and fourth region;
 - 5 i) the first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of the target gene;
 - ii) the second region comprising a nucleotide sequence of at least 19
10 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from the sense nucleotide region of the target gene ;
 - iii) the first region and said second region being capable of forming a double stranded RNA region:
 - 15 iv) the third region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene present in the eukaryotic cell and which is different from the target gene;
 - 20 v) the fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of the 19 consecutive nucleotides from the sense nucleotide region of the second gene;
 - vi) the third and fourth region being capable of forming a double
25 stranded RNA region; and
- c) a transcription termination and polyadenylation region that functions in cells of the eukaryotic organism, wherein said second gene is an endogenous gene of the eukaryotic organism or a transgene stably integrated into the genome of cells of the eukaryotic organism, as well
30 as the use of such a DNA molecule to measure the expression of a target gene by measuring the reduction in expression of a second gene.

The invention also provides non-human eukaryotic organisms comprising an RNA molecule or a DNA molecule as herein described.

5 Also provided by the invention is a method for identifying, within a population of dsRNA-mediated gene-silenced eukaryotic organisms, those organisms with the desired degree of silencing of a target gene comprising:

- a) providing the eukaryotic cells of the eukaryotic organisms with a dsRNA comprising a first region, a second region, a third region and a fourth region
10 region
 - i) the first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of the target gene;
 - 15 ii) the second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from the sense nucleotide region of the target gene ;
 - iii) the first region and the second region being capable of forming a double stranded RNA region:
20
 - iv) the third region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene present in the eukaryotic cells and which is different from the target gene;
25
 - v) the fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of the 19 consecutive nucleotides from the sense nucleotide region of the second gene;
 - 30 vi) the third and fourth region being capable of forming a double stranded RNA region; and

- b) identifying the organism with the desired degree of silencing of the target gene, by selecting those organisms with the desired degree of silencing of the second gene.

5 It is a further object of the invention to provide a method for modulating the reduction of the expression of a target gene in cells of a eukaryotic organism, comprising the steps of

- a) providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region,
 - 10 i) the first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of the target gene;
 - 15 ii) the second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the target gene;
 - iii) the first region and the second region being capable of forming a double stranded RNA region;
 - 20 iv) the third region and the fourth region comprising complementary nucleotide sequences which have a sequence identity of less than 50% to the nucleotide sequence of the target gene and which are capable of forming a double stranded RNA,
- 25 wherein the target gene is an endogenous gene in the eukaryotic cell or a transgene stably integrated in the genome of the eukaryotic cell. The size of the double stranded RNA capable of being formed by base-pairing between the third and fourth region may be equal or larger than the size of the double stranded RNA capable of being formed by base-pairing between the first and the second region.

The invention also provides dsRNA molecules suitable for the modulation of the degree of expression of a target gene in cells of eukaryotic organism, chimeric genes capable of yielding such dsRNA molecules, and non-human organisms containing such dsRNA molecules or chimeric genes.

5

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows seeds from T1 plants transformed with the FLC/CHS construct, fluorescing under UV light, compared to seeds from C24 wildtype (C24), a C24
10 wildtype containing a chimeric GUS gene under control of a CaMV35S promoter (GUS) and a homozygous CHS silenced line (CHS). Numbers indicate the days to flowering of the transgenic plant cell lines.

Figure 2 shows a Northern blot analysis of RNA prepared from transgenic plant
15 lines, comprising dsRNA encoding genes targeted towards both FLC and CHS, probed with an FLC probe (upper panel) and a CHS probe (lower panel) (Fig 2A). Fig 2B is a graphic representation of the the amount of mRNA detected from FLC expression plotted against the amount of mRNA detected from CHS expression.

20

DESCRIPTION OF DETAILED EMBODIMENTS

The invention is based on the unexpected observation that when a eukaryotic cell, such as a plant cell, comprises a double stranded RNA (dsRNA) wherein
25 the dsRNA comprises simultaneously complementary antisense and sense regions for at least two target genes, a correspondence exists between the degree of silencing of expression of all of the targeted genes. The correspondence further depends upon the relative size of the antisense and sense regions designed to reduce the expression of the different target genes.

30

This correspondence may conveniently be used to monitor silencing of genes, the expression of which results in a phenotype which is not straightforward to monitor. This may be accomplished by linking the complementary sense and antisense regions suitable to reduce or silence the expression of such a gene, to complementary sense and antisense regions suitable for reducing or silencing the expression of a gene whose expression results in a phenotype which can be monitored in a straightforward way.

Thus, in one embodiment of the invention, a method is provided for monitoring the reduction or silencing of the expression of a target gene in a eukaryotic cell, comprising the steps of

- Providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region, wherein
 - the first region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide region of the target gene;
 - the second region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from the sense nucleotide region of the target gene ;
 - the first region and second region are capable of forming a double stranded RNA region, preferably over the entire length of the first and second region, and at least over the length of the mentioned 19 nucleotides ;
 - the third region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide region of a second gene present in the eukaryotic cell and which is different from the target gene;

- the fourth region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide region of the same second gene;
- 5 - the third and fourth region are capable of forming a double stranded RNA region, preferably over the entire length of the first and second region, and at least over the length of the mentioned 19 nucleotides; and
- monitoring the expression of the degree of silencing of the target gene, by analyzing the reduction in expression of the second gene.

10

Preferably, the first and second region are capable of forming a double stranded RNA region at the same time as when the third and fourth region are forming a double stranded RNA.

- 15 The second gene different from the target gene is used as a reporter gene to monitor the degree of silencing of the target gene and whenever reference is made herein to a reporter gene, it is understood that this term is used to refer to a second gene present in the eukaryotic cell and which is different from the target gene.

20

For the purposes of the current invention, it is equal whether that second gene is an endogene, normally present in the eukaryotic cell, or a transgene, which has been introduced into the eukaryotic cell by human intervention at some point in history, preferably integrated into the genome of the eukaryotic cell in a

25 stable manner.

Preferably, the reporter gene should have a phenotype, the analysis of which is more straightforward than the analysis of the target gene or genes, either in terms of requirements of costs, expertise, time, labour, used apparatuses, etc.

30

Also preferably, the silencing of the reporter gene should not have a negative influence on the viability of the host cell or host organism, although the absence of negative influence may be dependent on particular conditions.

- 5 Conveniently, silencing the expression of the reporter gene may result in a visible phenotype, although the visibility of the phenotype may again be conditional.

Examples of reporter genes suitable for application of the methods of the invention in plant cells and plants include, but are not limited to Chalcone synthase gene (CHS, where down-regulation of the expression results in accumulation of UV-fluorescent compounds in the seed coat color); phytoene desaturase gene (PDS, where down-regulation of the gene expression results in photobleaching), flower locus C (FLC, where down-regulation of the gene expression results in early flowering), ethylene insensitivity gene 2 (EIN2, where down-regulation of the gene results in plants which are insensitive to ethylene, and will grow on media containing 1-aminocyclopropane-1-carboxylic acid(AAC)), visual marker gene such as seed coat color genes (e.g. R-gene in corn), plant-expressible GUS or GFP genes, phytochrome B and the like.

20

Examples of reporter genes suitable for application of the methods of the invention in animal cells and animals include but are not limited to GUS or GFP genes operably linked to expression regions suitable for animal cells but also genes such as *unc* in *Caenorhabditis elegans*, the silencing of which causes a characteristic twisting pattern in the nematodes, and the like.

25

Examples of reporter genes suitable for application of the methods of the invention in fungal cells include but are not limited to GUS or GFP genes operably linked to expression regions suitable for fungus cells but also genes the silencing of which causes auxotrophic growth (such as, e.g., *trpC*), a phenotype which can be easily screened on minimal media. It goes without

30

saying that in these cases, a master copy of the library of silenced fungal cells needs to be maintained under conditions allowing growth, *e.g.*, in the presence of the required nutrient compound.

- 5 Preferably, the reporter gene should be expressed at least under conditions where the target gene is expressed.

The 19 nucleotides which are at least 94% identical or complementary to a nucleotide sequence of 19 consecutive nucleotides of the target gene and/or
10 the reporter gene may be comprised within a larger RNA molecule. Such an RNA molecule will have a nucleotide sequence varying between as little as 19 bp to a length equal to the size of the target gene with a varying overall degree of sequence identity.

- 15 For the purpose of this invention, the "sequence identity" of two related nucleotide or amino acid sequences, expressed as a percentage, refers to the number of positions in the two optimally aligned sequences which have identical residues (x100) divided by the number of positions compared. A gap, *i.e.*, a position in an alignment where a residue is present in one sequence but
20 not in the other is regarded as a position with non-identical residues. The alignment of the two sequences is performed by the Needleman and Wunsch algorithm (Needleman and Wunsch 1970) The computer-assisted sequence alignment above, can be conveniently performed using standard software program such as GAP which is part of the Wisconsin Package Version 10.1
25 (Genetics Computer Group, Madison, Wisconsin, USA) using the default scoring matrix with a gap creation penalty of 50 and a gap extension penalty of 3. Sequences are indicated as "essentially similar" when such sequence have a sequence identity of at least about 75%, particularly at least about 80 %, more particularly at least about 85%, quite particularly about 90%, especially about
30 95%, more especially about 100%, quite especially are identical. It is clear than when RNA sequences are the to be essentially similar or have a certain degree

of sequence identity with DNA sequences, thymine (T) in the DNA sequence is considered equal to uracil (U) in the RNA sequence. Thus when it is stated in this application that a sequence of 19 consecutive nucleotides has a 94% sequence identity to a sequence of 19 nucleotides, this means that at least 18
5 of the 19 nucleotides of the first sequence are identical to 18 of the 19 nucleotides of the second sequence.

The mentioned sense or antisense nucleotide regions may thus be about 50 nt, 100nt, 200 nt, 300nt, 500nt, 1000 nt, 2000 nt or even about 5000 nt or larger in
10 length, each having an overall sequence identity of respectively about 40 %, 50%, 60 %, 70%, 80%, 90 % or 100%. The longer the sequence, the less stringent the requirement for the overall sequence identity.

The first, second, third and fourth region each may be separated by a spacer
15 region having a nucleotide sequence which is unrelated to the nucleotide sequence of either the target or the reporter gene.

For the purpose of the invention, although the regions are named consecutively , the order of the dsRNA regions in the dsRNA molecule is not important. In
20 other words, it does not matter whether, *e.g.*, the first or second region or alternatively the third or fourth region are located at the 5' or 3' end of the RNA molecule.

As used herein "comprising" is to be interpreted as specifying the presence of
25 the stated features, integers, steps or components as referred to, but does not preclude the presence or addition of one or more features, integers, steps or components, or groups thereof. Thus, *e.g.*, a nucleic acid or protein comprising a sequence of nucleotides or amino acids, may comprise more nucleotides or amino acids than the actually cited ones, *i.e.*, be embedded in a larger nucleic
30 acid or protein. A chimeric gene comprising a DNA region which is functionally or structurally defined, may comprise additional DNA regions etc.

An dsRNA molecule comprising a first, second, third and fourth region as herein defined may thus additionally include, *e.g.*, a fifth and sixth region having a nucleotide region of 19 nucleotides with at least 94% sequence identity or
5 complementarity to, *e.g.*, a target gene, which may be the same target gene as the first target gene, or may be a different one.

In one embodiment of the invention, the dsRNA molecule may further comprise one or more regions having at least 94% sequence identity to regions of 19
10 consecutive nucleotides from the sense nucleotide of the target genes, different from the 19 consecutive nucleotides as defined in the first region, and one or more regions having at least 94% sequence identity to 19 consecutive nucleotides from the complement of the sense nucleotide of the target gene, different from the 19 consecutive nucleotides as defined in the second region,
15 wherein these additional regions can basepair amongst themselves. Similarly, the dsRNA may also comprise additionally one or more regions having at least 94% sequence identity to regions of 19 consecutive nucleotides of a reporter gene different from the 19 nucleotides of the third region and one or more regions having at least 94% sequence identity to the complementary regions of
20 19 consecutive nucleotides of the reporter genes, wherein these additional regions are capable of base-pairing between themselves. Again, no particular order of the regions is required, and these regions may be dispersed among each other. Thus, *e.g.*, dsRNA regions directed towards silencing of the target gene may be alternated with dsRNA regions directed towards silencing of the
25 reporter gene, provided of course that basepairing between complementary RNA regions is still possible.

Conveniently, the dsRNA as described may be introduced into the host cell by introduction and possible integration of a chimeric gene, transcription of which
30 yields such a dsRNA. Thus the invention is also aimed at providing such a chimeric gene comprising

- a promoter or a promoter region which is capable of being expressed in cells of the eukaryotic organism of interest; operably linked to a DNA region which when transcribed yields a dsRNA molecule comprising a first region, a second region, a third region and a fourth region, wherein
 - 5 - the first region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of the target gene;
 - 10 - the second region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the target gene ;
 - 15 - the first region and second region are capable of forming a double stranded RNA region, preferably over the entire length of the first and second region and at least between the mentioned 19 nucleotides of the first and second region;
 - 20 - the third region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of a second gene present in the eukaryotic cell and which is different from the target gene;
 - 25 - the fourth region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the same second gene;
 - 30 - the third and fourth region are capable of forming a double stranded RNA region, preferably over the entire length of the first and second region, and at least over the length of the mentioned 19 nucleotides;
- and wherein the double stranded RNA regions formed between the first and second region and the double stranded RNA region formed between the third and fourth region are about equal in size; and

- a transcription termination and polyadenylation region suitable for the eukaryotic cell of choice.

As used herein, the term "promoter" denotes any DNA which is recognized and bound (directly or indirectly) by a DNA-dependent RNA-polymerase during initiation of transcription. A promoter includes the transcription initiation site, and binding sites for transcription initiation factors and RNA polymerase, and can comprise various other sites (e.g., enhancers), at which gene expression regulatory proteins may bind.

10

The term "regulatory region", as used herein, means any DNA, that is involved in driving transcription and controlling (*i.e.*, regulating) the timing and level of transcription of a given DNA sequence, such as a DNA coding for a protein or polypeptide. For example, a 5' regulatory region (or "promoter region") is a DNA sequence located upstream (*i.e.*, 5') of a coding sequence and which comprises the promoter and the 5'-untranslated leader sequence. A 3' regulatory region is a DNA sequence located downstream (*i.e.*, 3') of the coding sequence and which comprises suitable transcription termination (and/or regulation) signals, including one or more polyadenylation signals.

20

In one embodiment of the invention the promoter is a constitutive promoter. In another embodiment of the invention, the promoter activity is enhanced by external or internal stimuli (inducible promoter), such as but not limited to hormones, chemical compounds, mechanical impulses, abiotic or biotic stress conditions. The activity of the promoter may also regulated in a temporal or spatial manner (tissue-specific promoters; developmentally regulated promoters).

In a particular embodiment of the invention, the promoter is a plant-expressible promoter. As used herein, the term "plant-expressible promoter" means a DNA sequence that is capable of controlling (initiating) transcription in a plant cell.

30

This includes any promoter of plant origin, but also any promoter of non-plant origin which is capable of directing transcription in a plant cell, *i.e.*, certain promoters of viral or bacterial origin such as the CaMV35S (Hapster et al., 1988), the subterranean clover virus promoter No 4 or No 7 (WO9606932), or
5 T-DNA gene promoters but also tissue-specific or organ-specific promoters including but not limited to seed-specific promoters (e.g., WO89/03887), organ-primordia specific promoters (An et al., 1996), stem-specific promoters (Keller et al., 1988), leaf specific promoters (Hudspeth et al., 1989), mesophyll-specific promoters (such as the light-inducible Rubisco promoters), root-specific
10 promoters (Keller et al., 1989), tuber-specific promoters (Keil et al., 1989), vascular tissue specific promoters (Peleman et al., 1989), stamen-selective promoters (WO 89/10396, WO 92/13956), dehiscence zone specific promoters (WO 97/13865) and the like.

15 In another particular embodiment of the invention, the promoter is a fungus-expressible promoter. As used herein, the term "fungus-expressible promoter" means a DNA sequence which is capable of controlling (initiating) transcription in a fungal cell such as but not limited to the *A. nidulans trpC* gene promoter, or the *S. cerevisiae* GAL4 promoter that is capable of controlling (initiating)
20 transcription in a fungal cell.

In yet another particular embodiment of the invention, the promoter is a animal-expressible promoter. As used herein, the term "animal-expressible promoter" means a DNA sequence which is capable of controlling (initiating) transcription
25 in an animal cell and including but not limited to SV40 late and early promoters, cytomegalovirus CMV-IE promoters, RSV-LTR promoter, SCSV promoter, SCBV promoter and the like.

The dsRNA molecules useful for the invention may also be produced by *in vitro*
30 transcription. To this end, the promoter of the chimeric genes according to the invention may be a promoter recognized by a bacteriophage single subunit

RNA polymerase, such as the promoters recognized by bacteriophage single subunit RNA polymerase such as the RNA polymerases derived from the E. coli phages T7, T3, ϕ I, ϕ II, W31, H, Y, A1, 122, cro, C21, C22, and C2; *Pseudomonas putida* phage gh-1; *Salmonella typhimurium* phage SP6; *Serratia marcescens* phage IV; *Citrobacter* phage VIII; and *Klebsiella* phage No.11 [Hausmann, Current Topics in Microbiology and Immunology, 75: 77-109 (1976); Korsten et al., J. Gen Virol. 43: 57-73 (1975); Dunn et al., Nature New Biology, 230: 94-96 (1971); Towle et al., J. Biol. Chem. 250: 1723-1733 (1975); Butler and Chamberlin, J. Biol. Chem., 257: 5772-5778 (1982)]. Examples of such promoters are a T3 RNA polymerase specific promoter and a T7 RNA polymerase specific promoter, respectively. A T3 promoter to be used as a first promoter in the CIG can be any promoter of the T3 genes as described by McGraw et al, Nucl. Acid Res. 13: 6753-6766 (1985). Alternatively, a T3 promoter may be a T7 promoter which is modified at nucleotide positions -10, -11 and -12 in order to be recognized by T3 RNA polymerase [(Klement et al., J. Mol. Biol. 215, 21-29(1990))]. A preferred T3 promoter is the promoter having the "consensus" sequence for a T3 promoter, as described in US Patent 5,037,745. A T7 promoter which may be used according to the invention, in combination with T7 RNA polymerase, comprises a promoter of one of the T7 genes as described by Dunn and Studier, J. Mol. Biol. 166: 477-535 (1983). A preferred T7 promoter is the promoter having the "consensus" sequence for a T7 promoter, as described by Dunn and Studier (supra).

dsRNA can be produced in large amounts by contacting the acceptor vector DNA with the appropriate bacteriophage single subunit RNA polymerase under conditions well known to the skilled artisan. The so-produced dsRNA can then be used for delivery into cells prone to gene silencing, such as plant cells, fungal cells or animal cells. dsRNA may be introduced in animal cells via liposomes or other transfection agents (e.g. Clonfection transfection reagent or the CalPhos Mammalian transfection kit from ClonTech) and could be used for methods of treatment of animals, including humans, by silencing the

appropriate target genes. dsRNA can be introduced into the cell in a number of different ways. For example, the dsRNA may be administered by microinjection, bombardment by particles covered by the dsRNA, soaking the cell or organisms in a solution of the dsRNA, electroporation of cell membranes
5 in the presence of dsRNA, liposome mediated delivery of dsRNA and transfection mediated by chemicals such as calcium phosphate, viral infection, transformation and the like. The dsRNA may be introduced along with components that enhance RNA uptake by the cell, stabilize the annealed strands, or otherwise increase inhibition of the target gene. In the case of a
10 whole animal, the dsRNA is conveniently introduced by injection or perfusion into a cavity or interstitial space of an organism, or systemically via oral, topical, parenteral (including subcutaneous, intramuscular or intravenous administration), vaginal, rectal, intranasal, ophthalmic, or intraperitoneal administration. The dsRNA may also be administered via an implantable
15 extended release device.

The chimeric genes according to the invention capable of producing a dsRNA may also be equipped with any prokaryotic promoter suitable for expression of dsRNA in a particular prokaryotic host. The prokaryotic host can be used as a
20 source of dsRNA, e.g. by feeding it to an animal, such as a nematode or an insect, in which the silencing of the target gene is envisioned and monitored by reduction of the expression of the reporter gene. In this case, it will be clear that the target gene and reporter genes should be genes present in the cells of the target eukaryotic organism and not of the prokaryotic host organism. The
25 dsRNA according to the invention or chimeric genes capable of yielding such dsRNA molecules, can be thus produced in one host organism. The dsRNA can be administered to another target organisms (e.g. through feeding, orally administering, as a naked DNA or RNA molecule or encapsulated in a liposome, in a virus particle or attenuated virus particle, or on an inert particle etc.) and
30 effect reduction of gene expression in the target gene or genes and reporter gene or genes in another organism. In this case, it will be clear that the target

gene and reporter genes should be genes present in the cells of the target (eukaryotic) organism and not of the host organism.

Suitable transcription termination and polyadenylation regions include but are not limited to the SV40 polyadenylation signal, the HSV TK polyadenylation signal, the nopaline synthase gene terminator of *Agrobacterium tumefaciens*, the terminator of the CaMV 35S transcript, terminators of the subterranean stunt clover virus, the terminator of the *Aspergillus nidulans* trpC gene and the like.

10

The invention also aims at providing the dsRNA molecules, which may be obtained by transcription from these chimeric genes, and which are useful for the methods according to the invention.

15 It is another object of the invention to provide eukaryotic cells, and eukaryotic non-human organisms containing the dsRNA molecules of the invention, or containing the chimeric genes capable of producing the dsRNA molecules of the invention. In a preferred embodiment the chimeric genes are stably integrated in the genome of the cells of the eukaryotic organism.

20

In another preferred embodiment, the chimeric genes may be provided on a DNA molecule capable of autonomously replicating in the cells of the eukaryotic organism, such as e.g. viral vectors. The chimeric gene or the dsRNA may be also be provided transiently to the cells of the eukaryotic organism.

25

The dsRNA molecules according to this first aspect of the invention can also be used for identifying, within a population of dsRNA-mediated gene-silenced organisms, the ones with the desired degree of silencing of a target gene or genes. To this end, a population of dsRNA containing organisms is generated, wherein the dsRNA comprising a first region, a second region, a third region

30

and a fourth region, wherein the first region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of the target gene. The second region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the target gene. The third region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of a second gene present in the eukaryotic cell and which is different from the target gene. The fourth region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the same second gene. The first and second and the third and fourth region are capable of forming a double stranded RNA region, preferably over the entire length of the first and second or third and fourth region, and at least between the mentioned 19 nucleotide stretches, and wherein the double stranded RNA regions formed between the first and second region and the double stranded RNA region formed between the third and fourth region are about equal in size. The population of dsRNA-containing organisms is then analyzed for the down-regulation of the expression of the reporter gene and dsRNA-containing organisms wherein the down-regulation of the expression of the reporter gene corresponds to the desired degree of down-regulation are selected and isolated.

25

The correspondence in the degree of silencing of two or more genes induced by the presence of dsRNA that comprises the double stranded complementary sense and antisense RNA regions for the two or more target genes, may also be used to modulate the degree of silencing of those target genes. The correspondence in the degree of silencing of two or more genes is dependent on the relative size of the double stranded complementary sense and antisense

30

RNA regions. Consequently, the degree of silencing of one particular gene may be modulated by including into the same dsRNA molecule comprising complementary sense and antisense RNA regions, designed to silence the expression of a particular target gene or genes, an excess of unrelated complementary sequences, also capable of forming a double stranded RNA by base-pairing.

Without restricting the invention to a particular mode of action, it is thought that the enzyme in eukaryotic cells which is responsible for generating the small RNA molecules of about 21 nt from the dsRNA (such as DICER in *Drosophila*) may be saturated by including into the dsRNA an excess of dsRNA sequences (*i.e.*, complementary RNA molecules), the sequence of which is unrelated to the nucleotide sequence of the target gene or genes to be silenced.

Thus in one embodiment of the invention, a method is provided for modulating the reduction of the expression of a target gene in a eukaryotic organism, comprising the steps of

- providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region, wherein
 - the first region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of the target gene;
 - the second region comprises a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from the sense nucleotide sequence of the target gene ;
 - the first region and second region are capable of forming a double stranded RNA region preferably over the entire length of the first and second region, and at least between the mentioned stretches of 19 nucleotides;

- the third region and fourth region comprise complementary nucleotide sequences unrelated to the nucleotide sequence of the target gene and are capable of forming a double stranded RNA. The size of the double stranded RNA capable of being formed by base-pairing between the third and fourth region, is preferably equal or larger than the size of the double stranded RNA capable of being formed by base-pairing between the first and the second region. The size of the third/fourth region may also be smaller than the size of the first/second region, depending on the desired degree of modulation of the silencing.

10

Preferably, the first and second region are capable of forming a double stranded RNA region at the same time as when the third and fourth region are forming a double stranded RNA. In addition, preferably, the target gene is an endogenous gene of the eukaryotic cell or a transgene, stably integrated into the genome of the eukaryotic cells.

15

As used herein, a sequence is said to be "unrelated" to a target gene, when the overall sequence identity between the unrelated sequence and the target sequence is less than 50%, particularly less than 45%, more particularly less than 35%. The unrelated sequence preferably should not have a nucleotide sequence having at least 94% sequence identity with a stretch of 19 consecutive nucleotides of the target gene or with the complement thereof.

20

The natural variation in down regulation of the expression of a target gene occurring between different lines of a eukaryotic organism comprising the same dsRNA molecule will be shifted towards the lower end of the spectrum of gene silencing. This is due to inclusion of extra dsRNA nucleotide sequences unrelated to the target gene, which are operably linked to the dsRNA formed by the first and second RNA region.

25

30

The ratio of the size of the double stranded RNA capable of being formed by base-pairing between the third and fourth region to the size of the double stranded RNA capable of being formed by base-pairing between the first and the second region may vary from 0,1 to 20, preferably from 1 to 10. The greater
5 the mentioned ratio, the more organisms in a population of eukaryotic organisms comprising that dsRNA will have a lower degree of gene-silencing (*i.e.*, a higher level of expression of the target gene).

The 19 nucleotides of the first or second RNA region which are at least 94%
10 identical or complementary to a nucleotide sequence of 19 consecutive nucleotides of the target gene may be comprised within a larger RNA molecule. Such a longer RNA molecule will have a nucleotide sequence varying between as little as 19 bp to a length equal to the size of the target gene with a varying overall degree of sequence identity.

15

The mentioned sense or antisense nucleotide regions may thus be about 50 nt, 100nt, 200 nt, 300nt, 500nt, 1000 nt, 2000 nt or even about 5000 nt or larger in length, each having an overall sequence identity of respectively about 40 %, 50%, 60 %, 70%, 80%, 90 % or 100%. The longer the sequence, the less
20 stringent the requirement for the overall sequence identity.

The first, second, third and fourth region each may be separated by a spacer region having a nucleotide sequence which is unrelated to the nucleotide sequence of either the target or the reporter gene.

25

Conveniently, the dsRNA molecules may be introduced into the eukaryotic cells or organisms by transcription from a chimeric gene comprising a promoter or promoter region capable of transcribing a DNA region, which when transcribed yields the dsRNA according to this aspect of the invention. It is understood that
30 different embodiments concerning suitable promoters etc. mentioned in relation to the first aspect of the invention can also be applied here.

The invention also aims at providing dsRNA or chimeric genes capable of producing such dsRNA molecules according to the second aspect of the invention, as well as non-human eukaryotic organisms comprising such dsRNA
5 or chimeric genes.

The inverted repeat part of the chimeric genes of the invention can be conventionally constructed by recombinational cloning, using the means and methods described in US patent applications 60/244067 and 60/333743 (herein
10 incorporated by reference) or WO02/059294. Vectors according to US patent applications 60/244067 or US 60/333743 may be modified to include a third and fourth region capable of forming a double stranded RNA gene, wherein the third and fourth region are capable of downregulating the expression of a marker gene such as EIN2, PHYB, FLC and CHS.

15 The chimeric genes encoding dsRNA according to the invention preferably comprise an intron, preferably in the region between the second and third region, in order to increase the efficiency (as described in WO99/53050). As used herein, an "intron" or intervening sequence is used to refer to a DNA
20 region within a larger transcribed region, which is transcribed in the nucleus to yield an RNA region which is part of a larger RNA, however, said RNA corresponding to the intron sequence is removed from the larger RNA when transferred to the cytoplasm. The corresponding RNA is also referred to as an intron or intervening sequence. Intron sequences are flanked by splicing sites,
25 and synthetic introns may be made by joining appropriate splice sites to any sequence, having an appropriate branching point. Examples of introns include the pdk2 intron, catalase intron from Castor bean, Delta 12 desaturase intron from cotton, Delta 12 desaturase intron from *Arabidopsis*, ubiquitin intron from maize or the intron from SV40.

30

In one embodiment of the invention, combining the first and second aspect of the invention, a dsRNA molecule may be provided to a eukaryotic cell for monitoring the down regulation of the expression of a target gene by analyzing the down regulation of a second or reporter gene, as described herein. In such an embodiment, the ratio between the size of the dsRNA having sequence to the target gene or its complement and the size of the dsRNA having sequence to the second gene or its complement is between 2 and 10. It is expected that in this way, by identifying the organisms wherein the expression of second gene is effectively downregulated, a population of organisms wherein the expression of the target gene is severely downregulated will be identified.

The methods and means described herein, can be applied to any eukaryotic organism in which gene-silencing takes place, including but not limited to plants (such as corn, cotton, *Arabidopsis*, rice, vegetables, soybeans, tobacco, trees etc.) invertebrate animals (such as insects, shellfish, molluscs, crustaceans such as crabs, lobsters and prawns) vertebrate animals (fish, avian animals, mammals, humans), yeast and fungi amongst others.

The following non-limiting Examples describe method and means for monitoring dsRNA mediated silencing of the expression of a target gene through analysis of a second gene, as well as methods and means for modulating the degree of dsRNA mediated gene silencing in eukaryotic cells.

Unless stated otherwise in the Examples, all recombinant DNA techniques are carried out according to standard protocols as described in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, NY and in Volumes 1 and 2 of Ausubel *et al.* (1994) *Current Protocols in Molecular Biology*, *Current Protocols*, USA. Standard materials and methods for plant molecular work are described in *Plant Molecular Biology Labfax* (1993) by R.D.D. Croy, jointly published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications, UK. Other

references for standard molecular biology techniques include Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual*, Third Edition, Cold Spring Harbor Laboratory Press, NY, Volumes I and II of Brown (1998) *Molecular Biology LabFax*, Second Edition, Academic Press (UK). Standard
5 materials and methods for polymerase chain reactions can be found in Dieffenbach and Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, and in McPherson et al. (2000) *PCR - Basics: From Background to Bench*, First Edition, Springer Verlag, Germany.

10 Throughout the description and Examples, reference is made to the following sequences:

SEQ ID N° 1: oligonucleotide primer for PCR reaction of the CHS/FLC fragment.

15

SEQ ID N°2: oligonucleotide primer for PCR reaction of the CHS/FLC fragment.

SEQ ID N°3: oligonucleotide primer for PCR reaction of the CHS/FLC fragment.

20 SEQ ID N°4: oligonucleotide primer for PCR reaction of the CHS/FLC fragment.

SEQ ID N°5: attB1 recombination site.

SEQ ID N° 6: attB2 recombination site.

25

SEQ ID N°7: pHELLSGATE 8.

EXAMPLES

Example 1. An ihpRNA construct targeting two genes gives stoichiometric silencing.

5

Construction of an ihpRNA construct targeting two genes.

A chimeric gene encoding an ihpRNA construct targeting two distinct genes was constructed comprising 300 nt of the Flower Locus C (*FLC*) gene (Genbank accession AF116527, bases 620-920) followed by 300 nt of the Chalcone synthase (*CHS*) gene (Genbank accession Y18603, bases 147-532) in sense orientation, and a nucleotide sequence complementary to the mentioned 300 nt of CHS followed by the complementary nucleotide sequence to the mentioned 300 nt of FLC.

15

The CHS/FLC insert was generated by amplifying a 300 bp fragment of FLC with the oligonucleotide primers having the following sequences:

- 5'CTCGAGTCTAGAGGAGCAGAAGCTGAGATGGAG 3' (SEQ ID N° 1) (primer 138)
- 20 - 5'CTGCAGGAATAAGGTACAAAGTTCATC 3' (SEQ ID N° 2) (primer 136) and cloning the resulting product into pGEM T-easy (Promega, Madison, WI).

The CHS fragment was amplified with the oligonucleotide primers having the following sequences:

- 25 - 5' CTGCAGGCACTGCTAACCCTGAGAAAC 3' (SEQ ID N° 3) (primer 133)
- 5' GGTACCTTGACTTGGGCTGGCCCCACT 3' (SEQ ID N° 4) (primer 143) and also cloned into pGEM T-easy.

A 300 bp fragment was excised from pGEM T-easy CHS containing the CHS fragment and inserted into the PstI site of pGEM T-easy FLC. The resulting CHS/FLC chimeric insert was amplified using primers with the same gene-

30

specific sequence as 138 and 143 but modified to comprise also the nucleotide sequence encoding the recombination sites attB1 (GGGGACAAGTTTGTACAAAAAAGCAGGCT; SEQ ID N° 5) and attB2 (GGGACCACTTTGTACAAGAAAGCTGGGT; SEQ ID N° 6), respectively at
5 their 5' ends with F1 Taq DNA polymerase (Fisher Biotec, Subiaco, WA, Australia) using the manufacturer's protocol.

PCR products were precipitated by adding 3 volumes TE and two volumes 30% (w/v) PEG 3000, 30mM MgCl₂ and centrifuging at 13000 g for 15 minutes.
10 Recombination reaction of PCR products with pDONR201 (Invitrogen, Groningen, The Netherlands) were carried out in a total volume of 10 µL with 2 µL BP clonase buffer (Invitrogen), 1-2 µL PCR product 150 ng plasmid vector and 2 µL BP clonase (Invitrogen). The reaction was incubated at room temperature (25°C) for 1 h to overnight. After the incubation, 1 µL proteinase K
15 (2 µg/µL; Invitrogen) was added and incubated for 10 min at 37°C. 1-2 µL of the mix was used to transform E. coli DH5α, colonies were selected on the appropriate antibiotics. Clones were checked either by digestion of DNA minipreps or PCR.

20 Recombination reactions from pDONR201 clones to pHellsgate 8 were carried out in 10 µL total volume with 2 µL LR clonase buffer (Invitrogen), 2 µL pDONR201 clone (approximately 150 ng), 300 ng pHellsgate 8 and 2 µL LR clonase (Invitrogen). The reaction was incubated overnight at room temperature the proteinase treated and used to transform DH5α as for the BP
25 clonase reaction.

pHELLSGATE 8 is a vector suitable for recombinational cloning in such a way that an inverted repeat of the insert DNA of interest is generated. The construction of pHELLSGATE 8, and its use for recombination cloning have
30 been described in PCT application PCT/AU02/00073, US 60/264,067 or US

priority application US 60/333,743 (all documents herein incorporated by reference). pHELLSGATE 8 comprises the following DNA elements encoding

- a right border sequence of *Agrobacterium tumefaciens* T-DNA region;
- 5 • a CaMV 35S promoter region;
- an *attR1* recombination site;
- a *ccdB* selection marker gene from *E. coli*;
- an *attR2* recombination site;
- 10 • a *pdk2* (*Flavineria trinerva* pyruvate orthophosphate dikinase intron 2) intron sequence;
- a *attR2* recombination site;
- a *ccdB* selection marker gene;
- an *attR1* recombination site;
- 15 • a terminator region from *Agrobacterium tumefaciens* octopine synthase gene (*ocs*);
- a chimeric *nptII* gene flanked by a nopaline synthase gene (*nos*) promoter and *nos* terminator;
- a left border *Agrobacterium tumefaciens* T-DNA region;
- 20 • an origin of replication for *E. coli* and a streptomycin resistance gene.

The complete sequence of pHELLSGATE 8 is represented in SEQ ID N° 7.

Upon recombinational cloning of the CHS/FLC, PCR amplified insert as described above, a chimeric gene is thus generated comprising:

- 25 - a CaMV35S promoter region;
- a nucleotide sequence corresponding to bases 620-920 of *FLC*;
- a nucleotide sequence corresponding to bases 147-532 of *CHS*;
- a *pdk2* intron
- a nucleotide sequence complementary to bases 147-532 of *CHS*;
- 30 - a nucleotide sequence complementary to bases 620-920 of *FLC*;
- an *ocs* terminator region.

This chimeric gene is located between *Agrobacterium* T-DNA border sequences, together with a plant-expressible selectable marker gene. The vector can thus be introduced into *Agrobacterium* comprising the required helper functions, and used to transform plant cells.

Plant transformation.

Transformation of *Arabidopsis* C24 ecotype was via the floral dip method (Clough and Bent, 1998). Plants were selected on agar solidified MS media supplemented with 100 mg/l timentin and 50 mg/l kanamycin.

Phenotypic analysis methods for plants wherein FLC or CHS genes are silenced.

T1 *FLC* ihpRNA plants were scored by transferring to MS plates and scoring days to flower or rosette leaves at flowering compared to C24 wild type plants and *flc* mutant lines. Plant lines wherein the *FLC* gene expression is reduced or eliminated flower earlier than wild type C24 ecotype. A block in anthocyanin biosynthesis by reducing expression of chalcone synthase gene (*CHS*) leads to the accumulation of malonyl-CoA, which increases fluorescence of seed observed under UV light.

All of the T1 plants transformed with the iph *CHS/FLC* construct flowered earlier than wild type C24 (36 days) and later than the transposon-tagged *flc*-13 line (17days; Sheldon et al, 1999). The seeds from four iph *CHS/FLC* plants that flowered at different times were tested for fluorescence (Figure 1) compared to wildtype C24, a 35S::GUS line and a homozygous ihpRNA-silenced *CHS* line (Wesley et al, 2001). The relative concentrations of the different intermediates in the chalcone synthase pathway were also determined by HPLC analysis.

The CHS ihpRNA line shows strong fluorescence reflecting the accumulation of malonyl-CoA in this line while the wildtype C24 and the 35S::GUS lines do not show fluorescence under UV light. The four ihp CHS/FLC lines examined all show some fluorescence with the earliest flowering lines showing the strongest
5 fluorescence.

Example 2. Modulation of the degree of silencing of CHS gene by dilution with extra non-target sequences in the stem of the iphRNA.

10

To test the influence of the inclusion of extra nucleotides, unrelated to the target DNA sequences, in the ihpRNA or dsRNA molecules on the degree of silencing of a target gene, a number of chimeric genes were generated each comprising 100 bp from the CHS nucleotide sequence (corresponding to
15 Genbank accession Y18603) both in anti-sense and sense orientation (referred to as CHS_{antisense} and CHS_{sense} respectively.)

Progressively more nucleotide sequences (from 100 bp to 900 bp; referred to as Extra-NNN-nt_{sense} or Extra-NNN-nt_{antisense}) unrelated to the CHS gene were
20 inserted both in antisense and sense orientation. The unrelated nucleotide sequences were derived from the FLC gene (corresponding to Genbank accession Nr AF116527)

The different CHS/Extra nucleotide constructs were PCR amplified with attB1
25 and attB2 extended primers and introduced into pHELLSGATE 8 as described in Example 1. The resulting plasmids thus contain the following chimeric dsRNA genes:

1. <CaMV35S-CHS_{sense}-Extra_100_nt_{sense}-pdk2intron-Extra_100_nt_{antisense}-CHS_{antisense}-ocs terminator> or
- 30 2. <CaMV35S-CHS_{sense}-Extra_300_nt_{sense}-pdk2intron-Extra_300_nt_{antisense}-CHS_{antisense}-ocs terminator> or
3. <CaMV35S-CHS_{sense}-Extra_500_nt_{sense}-pdk2intron-Extra_500_nt_{antisense}-CHS_{antisense}-ocs terminator> or
4. <CaMV35S-CHS_{sense}-Extra_900_nt_{sense}-pdk2intron-Extra_900_nt_{antisense}-CHS_{antisense}-ocs terminator>

Transformed *Arabidopsis* plant lines comprising one of the four above mentioned chimeric genes were generated as described in Example 1, and fluorescence under UV light to monitor silencing of the CHS gene was scored as described in Example 1 for seeds of plants of each line. Independent transgenic lines were classified according to the degree of fluorescence. The results are summarized in Table 1.

Table 1: Number of independent transgenic lines showing different classes of fluorescence.

| | Fluorescence level | | | |
|--------------------------|--------------------|----|----|-----|
| | - | + | ++ | +++ |
| 100 nt CHS+ 100 extra nt | 12 | 0 | 20 | 0 |
| 100 nt CHS+ 300 extra nt | 11 | 18 | 11 | 1 |
| 100 nt CHS+ 500 extra nt | 5 | 7 | 13 | 0 |
| 100 nt CHS+ 900 extra nt | 18 | 8 | 3 | 0 |

From these results, it is clear that the higher the ratio of extra nucleotide sequences versus target sequences in the ihpRNA construct, the higher the proportion of transgenic plant lines obtained with a lower degree of fluorescence, *i.e.*, with a lower degree of silencing of the expression of the CHS target gene.

Inclusion of extra, unrelated, sequences in ihpRNA or dsRNA effectively shifts the distribution of silencing in the population of transgenic lines towards the lower end of the spectrum of gene-silencing.

Example 3. Northern analysis of transgenic plant lines comprising a chimeric gene encoding simultaneously dsRNA regions targeted towards FLC and CHS.

- 5 RNA was prepared from different *Arabidopsis* plant lines comprising two FLC-CHS hairpin constructs. One construct was a modified Hellsgate12 vector (see WO02/059294) with a FLC hairpin next to the intron with a CHS fragment inserted in the attR sites. The other construct was a modified Hellsgate12 with a CHS hairpin next to the intron with an FLC fragment inserted into the attR
- 10 sites. Two identical gels were run and blotted and probed with either FLC or CHS antisense RNA probes. The blots were RNase treated to ensure the signals were specific. The resulting autoradiogram is represented in Figure 2A. Estimates of the amount of hybridization for the different probes and lines are summarized in Table 2 and graphically represented in Fig 2B. There appears
- 15 be a good linear direct relation between the amount of FLC and CHS signals in each line indicating that where one gene is silenced the other gene is also silenced at the same degree.

Table 2. Estimates of hybridization signals in different plants using either FLC or CHS antisense probes.

| Plant | FLC mRNA | CHS mRNA |
|----------------------|----------|----------|
| HG12 2FLC + 2CHS #1 | 13867 | 15154 |
| HG12 2FLC + 2CHS #2 | 21962 | 26467 |
| HG12 2FLC + 2CHS #3 | 40434 | 40096 |
| HG12 2FLC + 2CHS#4 | 0 | 0 |
| HG12 2FLC + 2CHS #5 | 41122 | 38987 |
| HG12 2FLC + 2CHS #6 | 16609 | 23091 |
| HG12 2FLC + 2CHS #7 | 18417 | 3357 |
| HG12 2FLC + 2CHS #8 | 16463 | 17400 |
| HG12 2FLC + 2CHS # 9 | 47210 | 59680 |
| HG12 2FLC + 2CHS #10 | 10792 | 18879 |
| HG12 2CHS + 2FLC #1 | 3399 | 9694 |
| HG12 2CHS + 2FLC #2 | 3692 | 14317 |
| HG12 2CHS + 2FLC #3 | 1725 | 51004 |
| HG12 2CHS + 2FLC #4 | 4620 | 16067 |
| HG12 2CHS + 2FLC #5 | 21826 | 38643 |
| HG12 2CHS + 2FLC #7 | 13288 | 34043 |
| HG12 2CHS + 2FLC #8 | 6716 | 5949 |
| HG12 2CHS + 2FLC #9 | 3027 | 8036 |
| HG12 2CHS + 2FLC #10 | 1021 | 4677 |
| C24 wildtype | 23376 | 36279 |

References

- An et al., 1996 *The Plant Cell* 8, 15-30
- Butler and Chamberlin, 1982 *J. Biol. Chem.*, 257: 5772-5778
- 5 Dunn and Studier 1983 *J. Mol. Biol.* 166: 477-535
- Dunn et al. , 1971 *Nature New Biology*, 230: 94-96
- Fire *et al.*, 1998 *Nature* 391, 806-811
- Hamilton *et al.* 1998 *Plant J.* 15: 737-746
- Hapster et al., 1988 *Mol. Gen. Genet.* 212, 182-190
- 10 Hausmann, 1976 *Current Topics in Microbiology and Immunology*, 75: 77-109
- Hudspeth et al., 1989 *Plant Mol Biol* 12: 579-589
- Keil et al., 1989 *EMBO J.* 8: 1323-1330
- Keller et al., 1988 *EMBO J.* 7: 3625-3633
- Keller et al., 1989 *Genes Devel.* 3: 1639-1646
- 15 Klement et al., 1990 *J. Mol. Biol.* 215, 21-29
- Korsten et al., 1975 *J. Gen Virol.* 43: 57-73
- McGraw et al, 1985 *Nucl. Acid Res.* 13: 6753-6766
- Needleman and Wunsch 1970
- Peleman et al., 1989 *Gene* 84: 359-369
- 20 Smith et al., 2000 *Nature* 407: 319-320
- Towle et al., *J. Biol. Chem.* 250: 1723-1733 (1975)
- Waterhouse *et al.* 1998 *Proc. Natl. Acad. Sci. USA* 95: 13959-13964

SEQUENCE LISTING

<110> Commonwealth Scientific and Industrial Research Organization
5
<120> Methods and means for monitoring and modulating gene silencing

<130> COLINA-PCT

10 <150> US 60363852

<151> 2003-03-14

<160> 7
15
<170> PatentIn version 3.0

<210> 1
<211> 33
20 <212> DNA
<213> Artificial

<220>
<223> oligonucleotide primer for PCR amplification of CHS/FLC fragment
25
<400> 1
ctcgagtcta gaggagcaga agctgagatg gag 33

<210> 2
30 <211> 27
<212> DNA
<213> Artificial

<220>
35 <223> oligonucleotide primer for PCR amplification of CHS/FLC fragment

<400> 2
ctgcaggaat aaggtacaaa gtccatc 27

40 <210> 3
<211> 27
<212> DNA

<213> Artificial

<220>

<223> oligonucleotide primer for PCR amplification of CHS/FLC fragment

5

<400> 3

ctgcaggcac tgctaaccct gagaaac 27

<210> 4

10 <211> 27

<212> DNA

<213> Artificial

<220>

15 <223> oligonucleotide primer for PCR amplification of CHS/FLC fragment

<400> 4

ggtaccttga cttgggctgg cccact 27

<210> 5

20 <211> 29

<212> DNA

<213> Artificial

<220>

25 <223> attB1 recombination site

<400> 5

ggggacaagt ttgtacaaaa aagcaggct 29

30 <210> 6

<211> 27

<212> DNA

<213> Artificial

35 <220>

<223> attB2 recombination site

<400> 6

40 ggaccacttt gtacaagaaa gctgggt 27

<210> 7

<211> 17476

<212> DNA

<213> Artificial

<220>

5 <223> plasmid pHELLSGATE 8

<400> 7

| | | | | | | | |
|----|------------|------------|------------|-------------|-------------|-------------|------|
| | ggccgcacta | gtgatatccc | gcggccatgg | cgggccgggag | catgcgacgt | cgggcccaat | 60 |
| | tcgccctata | gtgagtcgta | ttacaattca | ctggccgctcg | ttttacaacg | tcgtgactgg | 120 |
| 10 | gaaaaccctg | gcgttaccca | acttaatcgc | cttgacgacac | atcccccttt | cgccagctgg | 180 |
| | cgtaatagcg | aagaggcccg | caccgatcgc | ccttcccaac | agttgcgcag | cctgaatggc | 240 |
| | gaatggaaat | tgtaaacggt | aatgggtttc | tggagtttaa | tgagctaagc | acatacgtca | 300 |
| | gaaaccatta | ttgcgcgttc | aaaagtcgcc | taaggctact | atcagctagc | aaatatttct | 360 |
| | tgtcaaaaat | gctccactga | cgttccataa | attccccctc | gtatccaatt | agagtctcat | 420 |
| 15 | attcactctc | aatccaaata | atctgcaatg | gcaattacct | tatccgcaac | ttctttacct | 480 |
| | atctccgccc | ggatccgggc | aggttctctc | gccgcttggg | tggagaggct | attcggctat | 540 |
| | gactgggcac | aacagacaat | cggtctgctc | gatgccgccg | tggtccgggt | gtcagcgagc | 600 |
| | gggcgcccgg | ttctttttgt | caagaccgac | ctgtccgggt | ccctgaatga | actgcaggac | 660 |
| | gaggcagcgc | ggctatcgtg | gctggccacg | acgggcgttc | cttgcgcagc | tgtgctcgac | 720 |
| 20 | gttgtcactg | aagcgggaag | ggactggctg | ctattggggc | aagtgccggg | gcaggatctc | 780 |
| | ctgtcatctc | accttgctcc | tgccgagaaa | gtatccatca | tggctgatgc | aatgcggcgg | 840 |
| | ctgcatacgc | ttgatccggc | tacctgccca | ttcgaccacc | aagcgaaaca | tcgcatacgag | 900 |
| | cgagcacgta | ctcggatgga | agccgggtct | gtcgatcagg | atgatctgga | cgaagagcat | 960 |
| | cagggggctc | cgccagccga | actgttcgcc | aggctcaagg | cgcgcgatgc | cgacggcgag | 1020 |
| 25 | gatctcgtcg | tgaccatggg | cgatgcctgc | ttgccgaata | tcatgggtgga | aaatggccgc | 1080 |
| | ttttctggat | tcatcgactg | tggccggctg | ggtgtggcgg | accgctatca | ggacatagcg | 1140 |
| | ttggctaccc | gtgatattgc | tgaagagctt | ggcggcgaaat | gggctgaccg | cttcctcgtg | 1200 |
| | ctttacggta | tcgccgctcc | cgattcgcag | cgcatacgct | tctatcgctt | tcttgacgag | 1260 |
| | ttcttctgag | cgggactctg | gggttcgaaa | tgaccgacca | agcgacgccc | aacctgccat | 1320 |
| 30 | cacgagattt | cgattccacc | gccgccttct | atgaaagggt | gggcttcgga | atcgttttcc | 1380 |
| | gggacgccgg | ctggatgata | ctccagcgcg | gggatctcat | gctggagtgc | ttcggcccacc | 1440 |
| | ccgatccaac | acttacgttt | gcaacgtcca | agagcaaata | gaccacgaac | gccggaaggc | 1500 |
| | tgccgcagcg | tgtggattgc | gtctcaattc | tctcttgtag | gaatgcaatg | atgaatatga | 1560 |
| | tactgactat | gaaactttga | gggaatactg | cctagcaccg | tcacctcata | acgtgcatca | 1620 |
| 35 | tgcatgccct | gacaacatgg | aacatcgcta | tttttctgaa | gaattatgct | cgttggagga | 1680 |
| | tgtcgcggca | attgcagcta | ttgccaacat | cgaactaccc | ctcacgcatg | cattcatcaa | 1740 |
| | tattattcat | gcggggaaag | gcaagattaa | tccaactggc | aaatcatcca | gcgtgattgg | 1800 |
| | taacttcagt | tccagcgact | tgattcgttt | tgggtgctacc | cacgttttca | ataaggacga | 1860 |
| | gatggtggag | taaagaagga | gtgcgtcgaa | gcagatcggt | caaacatttg | gcaataaagt | 1920 |
| 40 | ttcttaagat | tgaatcctgt | tgccgggtct | gcgatgatta | tcatataatt | tctgttgaa | 1980 |
| | tacgttaagc | atgtaataat | taacatgtaa | tgcatgacgt | tatttatgag | atgggttttt | 2040 |
| | atgattagag | tcccgcaatt | atacatttaa | tacgcgatag | aaaacaaaat | atagcgcgca | 2100 |
| | aactaggata | aattatcgcg | cgcggtgtca | tctatgttac | tagatcgaat | taattccagg | 2160 |

| | | | | | | | |
|----|-------------|------------|-------------|-------------|-------------|-------------|------|
| | cgggtgaagg | caatcagctg | ttgcccgctc | cactgggtgaa | aagaaaaacc | accccagtac | 2220 |
| | attaaaaacg | tccgcaatgt | gttattaagt | tgtctaagcg | tcaatttggt | tacaccacaa | 2280 |
| | tatatcctgc | caccagccag | ccaacagctc | cccgaccggc | agctcggcac | aaaatcacca | 2340 |
| | ctcgatacag | gcagcccatc | agtcggggac | ggcgtcagcg | ggagagccgt | tgtaaggcgg | 2400 |
| 5 | cagactttgc | tcatgttacc | gatgctattc | ggaagaacgg | caactaagct | gccgggtttg | 2460 |
| | aaacacggat | gatctcgcgg | agggtagcat | gttgattgta | acgatgacag | agcgttgctg | 2520 |
| | cctgtgatca | aatatcatct | ccctcgcaga | gatccgaatt | atcagccttc | ttattcattt | 2580 |
| | ctcgcttaac | cgtgacaggc | tgctgatctt | gagaactatg | ccgacataat | aggaaatcgc | 2640 |
| | tggataaagc | cgctgaggaa | gctgagtggc | gctatttctt | tagaagtga | cgttgacgat | 2700 |
| 10 | gtcgacggat | cttttccgct | gcataaccct | gcttcggggg | cattatagcg | attttttcgg | 2760 |
| | tatatccatc | ctttttcgca | cgatatacag | gattttgcc | aagggttcgt | gtagactttc | 2820 |
| | cttggtgat | ccaacggcgt | cagccgggca | ggataggtga | agtaggcccc | cccgcgagcg | 2880 |
| | gggtgttcct | cttcaactgc | ccttattcgc | acctggcggt | gctcaacggg | aatcctgctc | 2940 |
| | tgcgaggctg | gccggctacc | gccggcgtaa | cagatgaggg | caagcgggatg | gctgatgaaa | 3000 |
| 15 | ccaagccaac | caggggtgat | gctgccaact | tactgattta | gtgtatgatg | gtgtttttga | 3060 |
| | gggtgtccag | tggcttctgt | ttctatcagc | tgtccctcct | gttcagctac | tgacgggggtg | 3120 |
| | gtgcgtaacg | gcaaaagcac | cgccggacat | cagcgctatc | tctgctctca | ctgccgtaaa | 3180 |
| | acatggcaac | tgcagttcac | ttacaccgct | tctcaaccgg | gtacgcacca | gaaaaatcatt | 3240 |
| | gatatggcca | tgaatggcgt | tggatgccgg | gcaacagccc | gcattatggg | cgttggcctc | 3300 |
| 20 | aacacgattt | tacgtcactt | aaaaaactca | ggccgcagtc | ggtaacctcg | cgcatacagc | 3360 |
| | cgggcagtg | cgctcatcgc | tgcgcggaaa | tggacgaaca | gtggggctat | gtcggggcta | 3420 |
| | aatcgcgcca | gcgctggctg | ttttacgcgt | atgacagtct | ccggaagacg | gttgttgcgc | 3480 |
| | acgtattcgg | tgaacgcact | atggcgacgc | tggggcgctc | tatgagcctg | ctgtcaccct | 3540 |
| | ttgacgtggt | gatatggatg | acggatggct | ggccgctgta | tgaatcccg | ctgaagggaa | 3600 |
| 25 | agctgcacgt | aatcagcaag | cgatatacgc | agcgaattga | gcggcataac | ctgaatctga | 3660 |
| | ggcagcacct | ggcacggcgt | ggacgggaagt | cgctgtcggt | ctcaaaatcg | gtggagctgc | 3720 |
| | atgacaaaagt | catcgggcat | tatctgaaca | taaaacacta | tcaataagtt | ggagtcatta | 3780 |
| | cccaaccagg | aagggcagcc | cacctatcaa | ggtgtactgc | cttccagacg | aacgaagagc | 3840 |
| | gattgaggaa | aaggcggcgg | cgccgggcat | gagcctgtcg | gcctacctgc | tggccgtcgg | 3900 |
| 30 | ccagggctac | aaaatcacgg | gcgtcgtgga | ctatgagcac | gtccgcgagc | tggcccgcgt | 3960 |
| | caatggcgac | ctgggcccgc | tggggggcct | gctgaaactc | tggctcaccg | acgacccgcg | 4020 |
| | cacggcgcg | ttcggtgatg | ccacgatcct | cgccctgctg | gcgaagatcg | aagagaagca | 4080 |
| | ggacgagctt | ggcaaggcca | tgatgggcgt | ggccgcggcg | agggcagagc | catgactttt | 4140 |
| | ttagccgcta | aaacggccgg | ggggtgcgcg | tgattgcca | gcacgtcccc | atgcgctcca | 4200 |
| 35 | tcaagaagag | cgacttcgcg | gagctgggat | tcgtgcaggg | caagattcgg | aataccaagt | 4260 |
| | acgagaagga | cggccagacg | gtctacggga | ccgacttcat | tgccgataag | gtggattatc | 4320 |
| | tggacaccaa | ggcaccaggc | gggtcaaata | aggaataagg | gcacattgcc | ccggcgtagg | 4380 |
| | tcggggcaat | cccgcgaagg | gggtgaatga | atcggaagct | tgaccggaag | gcatacaggc | 4440 |
| | aagaactgat | cgacgcgggg | ttttccgccc | aggatgccga | aaccatcgca | agccgcaccg | 4500 |
| 40 | tcatgcgtgc | gccccgcgaa | accttccagt | ccgtcggctc | gatggtccag | caagctacgg | 4560 |
| | ccaagatcga | gcgcgacagc | gtgcaactgg | ctccccctgc | cctgcccgcg | ccatcgcccg | 4620 |
| | ccgtggagcg | ttcgcgctcg | ctcgaacagg | aggcggcagg | tttggcgaag | tcgatgacca | 4680 |
| | tcgacacgcg | aggaactatg | acgaccaaga | agcgaaaaac | cgccggcgag | gacctggcaa | 4740 |

| | | | | | | | |
|----|-------------|------------|-------------|-------------|-------------|-------------|------|
| | aacagggtcag | cgaggccaag | caggccgcgt | tgctgaaaca | cacgaagcag | cagatcaagg | 4800 |
| | aaatgcagct | ttccttggtc | gatattgcgc | cgtggccgga | cacgatgcga | gcgatgccaa | 4860 |
| | acgacacggc | ccgctctgcc | ctgttcacca | cgcgcaacaa | gaaaatcccc | cgcgagggcg | 4920 |
| | tgcaaaacaa | ggtcattttc | cacgtcaaca | aggacgtgaa | gatcacctac | accggcgctcg | 4980 |
| 5 | agctgcgggc | cgacgatgac | gaactgggtg | ggcagcaggt | gttggagtag | gcgaagcgca | 5040 |
| | cccctatcgg | cgagccgatc | accttcacgt | tctacgagct | ttgccaggac | ctgggctggg | 5100 |
| | cgatcaatgg | ccggtattac | acgaaggccg | aggaatgcct | gtcgcgccta | caggcgacgg | 5160 |
| | cgatgggctt | cacgtccgac | cgcgttgggc | acctggaatc | ggtgtcgctg | ctgcaccgct | 5220 |
| | tccgcgtcct | ggaccgtggc | aagaaaacgt | cccgttgcca | ggtcctgata | gacgaggaaa | 5280 |
| 10 | tcgtcgtgct | gtttgctggc | gaccactaca | cgaaattcat | atgggagaa | taccgcaagc | 5340 |
| | tgctgcggac | ggcccgcgag | atgttcgact | atttcagctc | gcaccgggag | ccgtaccgcg | 5400 |
| | tcaagctgga | aaccttccgc | ctcatgtgct | gatcggattc | caccgcgctg | aagaagtggc | 5460 |
| | gcgagcaggt | cggcgaagcc | tgcaagaggt | tgccagggcag | cggcctgggtg | gaacacgcct | 5520 |
| | gggtcaatga | tgacctgggt | cattgcaaac | gctagggcct | tgtgggggtca | gttccggctg | 5580 |
| 15 | gggggttcagc | agccagcgct | ttactggcat | ttcaggaaca | agcgggcact | gctcgacgca | 5640 |
| | cttgcttcgc | tcagtatcgc | tcgggacgca | cggcgcgctc | tacgaactgc | cgataaacag | 5700 |
| | aggattaaaa | ttgacaattg | tgattaaagg | tcagattcga | cggcttggag | cggccgacgt | 5760 |
| | gcaggatttc | cgcgagatcc | gattgtcggc | cctgaagaaa | gctccagaga | tgttcgggtc | 5820 |
| | cgttttacgag | cacgaggaga | aaaagcccat | ggaggcgctt | gctgaacggg | tgcgagatgc | 5880 |
| 20 | cgtggcattc | ggcgcctaca | tcgacggcga | gatcattggg | ctgtcgggtc | tcaaacagga | 5940 |
| | ggacggcccc | aaggacgctc | acaaggcgca | tctgtccggc | gttttcgtgg | agcccgaaaca | 6000 |
| | gcgaggccga | ggggtcgccg | gtatgctgct | gcgggcggtt | ccggcggggt | tattgctcgt | 6060 |
| | gatgatcgtc | cgacagattc | caacgggaat | ctggtggatg | cgcattctca | tcctcggcgc | 6120 |
| | acttaatat | tcgctattct | ggagcttggt | gtttatttct | gtctaccgcc | tgccgggcgg | 6180 |
| 25 | ggtcgcggcg | acggtaggag | ctgtgcagcc | gctgatgggt | gtgttcatct | ctgccgctct | 6240 |
| | gctaggtagc | ccgatacgat | tgatggcggt | cctgggggct | atttgcgga | ctgcgggcgt | 6300 |
| | ggcgctgttg | gtgttgacac | caaacgcagc | gctagatcct | gtcggcgctc | cagcgggcct | 6360 |
| | ggcgggggcg | gtttccatgg | cggttcggaac | cgtgctgacc | cgcaagtggc | aacctcccgt | 6420 |
| | gcctctgctc | acctttaccg | cctggcaact | ggcggccgga | ggacttctgc | tcgttccagt | 6480 |
| 30 | agcttttagtg | tttgatccgc | caatcccgat | gcctacagga | accaatgttc | tcggcctggc | 6540 |
| | gtggctcggc | ctgatcggag | cgggtttaac | ctacttcctt | tggttccggg | ggatctcgcg | 6600 |
| | actcgaacct | acagttgttt | ccttactggg | ctttctcagc | cgggatggcg | ctaagaagct | 6660 |
| | attgccgccc | atcttcatat | gcggtgtgaa | ataccgcaca | gatgcgtaag | gagaaaatac | 6720 |
| | cgcatacggc | gctcttccgc | ttcctcgtct | actgactcgc | tgcgctcggg | cgttcggctg | 6780 |
| 35 | cggcgagcgg | tatcagctca | ctcaaaggcg | gtaatacggg | tatccacaga | atcaggggat | 6840 |
| | aacgcaggaa | agaacatgtg | agcaaaaggc | cagcaaaagg | ccaggaaccg | taaaaaggcc | 6900 |
| | gcgttgctgg | cgtttttcca | taggctccgc | ccccctgacg | agcatcaca | aatcgcagc | 6960 |
| | tcaagtacga | gggtggcga | cccgcagagga | ctataaagat | accaggcggt | tccccctgga | 7020 |
| | agctccctcg | tgcgctctcc | tggtccgacc | ctgccgctta | ccggataacct | gtccgccttt | 7080 |
| 40 | ctcccttcgg | gaagcgtggc | gctttctcaa | tgctcacgct | gtaggtatct | cagttcgggtg | 7140 |
| | taggtcgttc | gctccaagct | gggctgtgtg | cacgaacccc | ccgttcagcc | cgaccgctgc | 7200 |
| | gccttatccg | gtaactatcg | tcttgagtcc | aacccggtaa | gacacgactt | atcgccactg | 7260 |
| | gcagcagcca | ctggtaacag | gattagcaga | gcgaggatat | taggcgggtg | tacagagttc | 7320 |

| | | | | | | | |
|----|-------------|------------|-------------|------------|-------------|------------|------|
| | ttgaagtgg | ggcctaacta | cggtacact | agaaggacag | tatttggtat | ctgcgctctg | 7380 |
| | ctgaagccag | ttaccttcgg | aaaaagagtt | ggtagctctt | gatccggcaa | acaaaccacc | 7440 |
| | gctggtagcg | gtgggttttt | tgtttgcaag | cagcagatta | cgcgagaaa | aaaaggatat | 7500 |
| | caagaagatc | ctttgatctt | ttctacgggg | tctgacgctc | agtggaaacga | aaactcacgt | 7560 |
| 5 | taagggattt | tggtcatgag | attatcaaaa | aggatcttca | cctagatcct | tttaaattaa | 7620 |
| | aaatgaagtt | ttaaatcaat | ctaaagtata | tatgagtaaa | cttgggtctga | cagttaccaa | 7680 |
| | tgcttaatca | gtgaggcacc | tatctcagcg | atctgtctat | ttcgttcate | catagttgcc | 7740 |
| | tgactccccg | tcgtgtagat | aactacgata | cgggagggct | taccatctgg | ccccagtgtc | 7800 |
| | gcaatgatac | cgcgagaccc | acgctcaccg | gctccagatt | tatcagcaat | aaaccagcca | 7860 |
| 10 | gccggaaggg | ccgagcgag | aagtggctct | gcaactttat | ccgcctccat | ccagtctatt | 7920 |
| | aaacaagtgg | cagcaacgga | ttcgcaaaacc | tgctacgcct | tttgtgccaa | aagccgcgcc | 7980 |
| | aggtttgca | tccgctgtgc | caggcgcttag | gcgtcatatg | aagatttcgg | tgatccctga | 8040 |
| | gcaggtggcg | gaaacattgg | atgctgagaa | ccatttcatt | gttcgtgaag | tgctcgatgt | 8100 |
| | gcacctatcc | gaccaaggct | ttgaactatc | taccagaagt | gtgagccctt | accggaagga | 8160 |
| 15 | ttacatctcg | gatgatgact | ctgatgaaga | ctctgcttgc | tatggcgcat | tcacgacca | 8220 |
| | agagcttgtc | gggaagattg | aactcaactc | aacatggaac | gatctagcct | ctatcgaaca | 8280 |
| | cattgtttgt | tcgcacacgc | accgaggcaa | aggagtgcgc | cacagtctca | tcgaatttgc | 8340 |
| | gaaaaagtgg | gactaagca | gacagctcct | tgccatatac | ttagagacac | aaacgaacaa | 8400 |
| | tgtacctgcc | tgcaatttgt | acgcaaaatg | tggctttact | ctcggcgcca | ttgacctgtt | 8460 |
| 20 | cacgtataaa | actagacctc | aagtctcgaa | cgaaacagcg | atgtactggg | actggttctc | 8520 |
| | gggagcacag | gatgacgcct | aacaattcat | tcaagccgac | accgcttcgc | ggcgcggtt | 8580 |
| | aattcaggag | ttaaacatca | tgagggaagc | ggtgatcgcc | gaagtatcga | ctcaactatc | 8640 |
| | agaggtagtt | ggcgtcatcg | agcgccatct | cgaaccgacg | ttgctggccg | tacatttgta | 8700 |
| | cggctccgca | gtggatggcg | gcctgaagcc | acacagtgat | attgatttgc | tggttacggt | 8760 |
| 25 | gaccgtaagg | cttgatgaaa | caacgcggcg | agctttgatc | aacgaccttt | tggaaacttc | 8820 |
| | ggcttcccct | ggagagagcg | agattctccg | cgctgtagaa | gtcaccattg | ttgtgcacga | 8880 |
| | cgacatcatt | ccgtggcggt | atccagctaa | gcgcgaactg | caatttggag | aatggcagcg | 8940 |
| | caatgacatt | cttgcaggta | tcttcagacc | agccacgata | gacattgatc | tggctatctt | 9000 |
| | gctgacaaaa | gcaagagaac | atagcggttg | cttggtaggt | ccagcggcgg | aggaactctt | 9060 |
| 30 | tgatccggtt | cctgaacagg | atctatttga | ggcgctaaat | gaaaccttaa | cgctatggaa | 9120 |
| | ctcgccgccc | gactgggctg | gcgatgagcg | aaatgtagtg | cttacgttgt | cccgcatttg | 9180 |
| | gtacagcgca | gtaaccggca | aaatcgcgcc | gaaggatgtc | gctgccgact | gggcaatgga | 9240 |
| | gcgcctgccg | gcccagtatc | agcccgctcat | acttgaagct | aggcaggctt | atcttgga | 9300 |
| | agaagatcgc | ttggcctcgc | gcgcagatca | gttggagaaa | tttgttctact | acgtgaaagg | 9360 |
| 35 | cgagatcacc | aaggtagtcg | gcaataaatg | tctaacaatt | cgttcaagcc | gacgccgctt | 9420 |
| | cgcgccgcgg | cttaactcaa | gcgttagaga | gctggggaag | actatgcgcg | atctgttgaa | 9480 |
| | ggtggttcta | agcctcgtac | ttgcgatggc | atcggggcag | gcacttgctg | acctgccaat | 9540 |
| | tgtttttagtg | gatgaagctc | gtcttcccta | tgactactcc | ccatccaact | acgacatttc | 9600 |
| | tccaagcaac | tacgacaact | ccataagcaa | ttacgacaat | agtccatcaa | attacgacaa | 9660 |
| 40 | ctctgagagc | aactacgata | atagttcatc | caattacgac | aatagtcgca | acggaaatcg | 9720 |
| | taggcttata | tatagcgcaa | atgggtctcg | cactttcgcc | ggctactacg | tcattgccaa | 9780 |
| | caatgggaca | acgaacttct | tttccacatc | tggcaaaagg | atgttctaca | ccccaaaagg | 9840 |
| | ggggcgcgcc | gtctatggcg | gcaaagatgg | gagcttctgc | ggggcattgg | tcgtcataaa | 9900 |

| | | | | | | | |
|----|------------|------------|------------|-------------|------------|------------|-------|
| | tgcccaattt | tcgcttgccc | tgacagataa | cggcctgaag | atcatgtatc | taagcaacta | 9960 |
| | gcctgctctc | taataaaatg | ttaggagctt | ggctgccatt | tttggggtga | ggccgttcgc | 10020 |
| | ggccgagggg | cgcagccctt | ggggggatgg | gaggcccgcg | ttagcggggc | gggaggggtc | 10080 |
| | gagaaggggg | ggcacccttc | ttcggcgctg | gcgggtcacg | gccagggcgc | agccctgggt | 10140 |
| 5 | aaaaacaagg | tttataaata | ttggttttaa | agcagggttaa | aagacagggt | agcgggtggc | 10200 |
| | gaaaaacggg | cggaaccctt | tgcaaatgct | ggattttctg | cctgtggaca | gcccctcaaa | 10260 |
| | tgtcaatagg | tgcccccctc | atctgtcagc | actctgcccc | tcaagtgtca | aggatcgcg | 10320 |
| | ccctcatctg | tcagtagtcg | cgccctctca | gtgtcaatac | cgcagggcac | ttatccccag | 10380 |
| | gcttgctcac | atcatctgtg | ggaaactcgc | gtaaaatcag | gcgttttcgc | cgatttgcca | 10440 |
| 10 | ggctggccag | ctccacgtcg | ccggccgaaa | tcgagcctgc | ccctcatctg | tcaacgccgc | 10500 |
| | gccgggtgag | tcggccccct | aagtgtcaac | gtccgcccc | catctgtcag | tgagggccaa | 10560 |
| | gttttcgcgc | aggtatccac | aacgccggcg | gccggccgcg | gtgtctcgca | cacggcttcg | 10620 |
| | acggcggttc | tgccgcgttt | gcagggccat | agacggccgc | cagcccagcg | gcgagggcaa | 10680 |
| | ccagcccggg | gagcgtcgga | aagggtcgac | atcttgctgc | gttcggatat | tttcgtggag | 10740 |
| 15 | ttcccgcac | agaccgggat | tgaaggcgag | atccagcaac | tcgcgccaga | tcctcctgtg | 10800 |
| | acggaacttt | ggcgctgat | gactggccag | gacgtcggcc | gaaagagcga | caagcagatc | 10860 |
| | acgattttcg | acagcgtcgg | atctgcgac | gaggattttt | cggcgctgcg | ctacgtccgc | 10920 |
| | gaccgcgttg | agggatcaag | ccacagcagc | ccactcgacc | ttctagccga | cccagacgag | 10980 |
| | ccaagggatc | tttttggaat | gctgctccgt | cgtcaggctt | tccgacgttt | gggtggttga | 11040 |
| 20 | acagaagtca | ttatcgtacg | gaatgccagc | actcccagg | ggaaccctgt | ggttggcatg | 11100 |
| | cacatacaaa | tggaacgaac | gataaacctt | ttcacgccct | tttaaatatc | cgttattcta | 11160 |
| | ataaacgctc | ttttctctta | ggtttaccgc | ccaatatatc | ctgtcaaaca | ctgatagttt | 11220 |
| | aaactgaagg | cggaacacga | caatctgac | atgagcggag | aattaaggga | gtcacgttat | 11280 |
| | gacccccgcc | gatgacgcgg | gacaagccgt | tttacgtttg | gaactgacag | aaccgcaacg | 11340 |
| 25 | attgaaggag | ccactcagcc | ccaatacgca | aaccgcctct | ccccgcgcgt | tggccgattc | 11400 |
| | attaatgcag | ctggcacgac | aggtttcccg | actggaaagc | gggcagtgag | cgcaacgcaa | 11460 |
| | ttaatgtgag | ttagctcact | cattaggcac | cccaggcttt | acactttatg | cttcgggtc | 11520 |
| | gtatgttggt | tggaattgtg | agcggataac | aatttcacac | aggaaacagc | tatgaccatg | 11580 |
| | attacgcaa | gctatttagg | tgacactata | gaatactcaa | gctatgcac | caacgcgttg | 11640 |
| 30 | ggagctctcc | catatcgacc | tgacggcgcc | cgctcgacga | attaattcca | atcccacaaa | 11700 |
| | aatctgagct | taacagcaca | gttgctcctc | tcagagcaga | atcgggtatt | caacaccctc | 11760 |
| | atatcaacta | ctacgttggt | tataacggct | cacatgccgg | tatatagcat | gactgggggt | 11820 |
| | gtacaaaggc | ggcaacaaac | ggcgctcccg | gagttgcaca | caagaaattt | gccactatta | 11880 |
| | cagaggcaag | agcagcagct | gacgcgtaca | caacaagtca | gcaaacagac | aggttgaact | 11940 |
| 35 | tcatacccaa | aggagaagct | caactcaagc | ccaagagctt | tgctaaggcc | ctaacaagcc | 12000 |
| | caccaaagca | aaaagccac | tggttcacgc | taggaaccaa | aaggcccagc | agtgatccag | 12060 |
| | ccccaaaaga | gatctccttt | gccccggaga | ttacaatgga | cgatttcctc | tatctttacg | 12120 |
| | atctaggaag | gaagttcgaa | ggtgaagggt | acgacactat | gttcaccact | gataatgaga | 12180 |
| | aggttagcct | cttcaatttc | agaaagaatg | ctgaccacac | gatgggtaga | gaggcctacg | 12240 |
| 40 | cagcagggtc | catcaagacg | atctacccga | gtaacaatct | ccaggagatc | aaataccttc | 12300 |
| | ccaagaagg | taaagatgca | gtcaaaagat | tcaggactaa | ttgcatcaag | aacacagaga | 12360 |
| | aagacatatt | tctcaagatc | agaagtacta | ttccagtatg | gacgattcaa | ggcttgcttc | 12420 |
| | ataaaccaag | gcaagtaata | gagattggag | tctctaaaaa | ggtagttcct | actgaatcta | 12480 |

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| | aggccatgca | tggagtctaa | gattcaaate | gaggatctaa | cagaactcgc | cgtgaagact | 12540 |
| | ggcgaacagt | tcatacagag | tctttttacga | ctcaatgaca | agaagaaaat | cttcgtcaac | 12600 |
| | atggtggagc | acgacactct | ggtctactcc | aaaaatgtca | aagatacagt | ctcagaagac | 12660 |
| | caaagggcta | ttgagacttt | tcaacaaagg | ataatttcgg | gaaacctcct | cggattccat | 12720 |
| 5 | tgcccagcta | tctgtcactt | catcgaaagg | acagtagaaa | aggaagggtgg | ctcctacaaa | 12780 |
| | tgccatcatt | gcgataaagg | aaaggctatc | attcaagatc | tctctgccga | cagtgggtccc | 12840 |
| | aaagatggac | ccccacccac | gaggagcatc | gtggaaaaag | aagacgttcc | aaccacgtct | 12900 |
| | tcaaagcaag | tggattgatg | tgacatctcc | actgacgtaa | gggatgacgc | acaatcccac | 12960 |
| | tatccttcgc | aagacccttc | ctctatataa | ggaagttcat | ttcatttgga | gaggacacgc | 13020 |
| 10 | tcgagacaag | tttgtacaaa | aaagctgaac | gagaaacgta | aaatgatata | aatatcaata | 13080 |
| | tattaaatta | gatttttgcat | aaaaaacaga | ctacataata | ctgtaaaaca | caacatatcc | 13140 |
| | agtcactatg | aatcaactac | ttagatggta | ttagtgacct | gtagtgcacc | gacagccttc | 13200 |
| | caaagtgtct | tcgggtgatg | ctgccaaact | agtcgaccga | cagccttcca | aatgttcttc | 13260 |
| | tcaaacggaa | tcgtcgtatc | cagcctactc | gctattgtcc | tcaatgccgt | attaaatcat | 13320 |
| 15 | aaaaagaaat | aagaaaaaga | ggtgcgagcc | tcttttttgt | gtgacaaaat | aaaaacatct | 13380 |
| | acctattcat | atacgctagt | gtcatagtcc | tgaaaatcat | ctgcatcaag | aacaatttca | 13440 |
| | caactcttat | acttttctct | tacaagtcgt | tcggcttcat | ctggattttc | agcctctata | 13500 |
| | cttactaaac | gtgataaagt | ttctgttaatt | tctactgtat | cgacctgcag | actgggtgtg | 13560 |
| | tataagggag | cctgacattt | atattcccca | gaacatcagg | ttaatggcgt | ttttgatgtc | 13620 |
| 20 | atttttcgcg | tggctgagat | cagccacttc | ttccccgata | acggagaccg | gcacactggc | 13680 |
| | catatcgggtg | gtcatcatgc | gccagctttc | atccccgata | tgaccaccgc | ggtaaagtcc | 13740 |
| | acgggagact | ttatctgaca | gcagacgtgc | actggccagg | gggatcacca | tccgtcgccc | 13800 |
| | gggcgtgtca | ataatatcac | tctgtacatc | cacaaacaga | cgataacggc | tctctctttt | 13860 |
| | ataggtgtaa | accttaaact | gcatttcacc | agtcctgtgt | ctcgtcagca | aaagagccgt | 13920 |
| 25 | tcattttcaat | aaaccgggcg | acctcagcca | tcccttcctg | attttccgct | ttccagcgtt | 13980 |
| | cggcacgcag | acgacgggct | tcattctgca | tggttgtgct | taccagaccg | gagatatatga | 14040 |
| | catcatatat | gccttgagca | actgatagct | gtcgtgtgca | actgtcactg | taatacgtctg | 14100 |
| | cttcatagca | cacctctttt | tgacataact | cgggtagtgc | cgatcaacgt | ctcattttcg | 14160 |
| | ccaaaagttg | gcccagggct | tcccgggtatc | aacagggaca | ccaggattta | tttattctgc | 14220 |
| 30 | gaagtgatct | tccgtcacag | gtattttattc | ggcgcaaagt | gcgtcgggtg | atgctgccaa | 14280 |
| | cttagtcgac | tacaggtcac | taataccatc | taagtagttg | attcatagtg | actggatatg | 14340 |
| | ttgtgtttta | cagtattatg | tagtctgttt | tttatgcaaa | atctaattta | atatattgat | 14400 |
| | atttatatca | ttttacgttt | ctcgttcagc | tttctgttac | aaagtgggtc | cgagggaattc | 14460 |
| | gggtacccag | cttggttaagg | aaataattat | tttctttttt | ccttttagta | taaaatagtt | 14520 |
| 35 | aagtgatgtt | aattagtatg | attataataa | tatagttgtt | ataattgtga | aaaaataatt | 14580 |
| | tataaatata | ttgtttacat | aaacaacata | gtaatgtaaa | aaaatatgac | aagtgatgtg | 14640 |
| | taagacgaag | aagataaaaag | ttgagagtaa | gtatattatt | tttaatgaat | ttgatcgaac | 14700 |
| | atgtaagatg | atatactagc | attaatattt | gttttaataca | taatagtaat | tctagctggt | 14760 |
| | ttgatgaatt | aaatatcaat | gataaaatac | tatagtaaaa | ataagaataa | ataaattaaa | 14820 |
| 40 | ataatatttt | tttatgatta | atagttttatt | atataattaa | atatctatac | cattactaaa | 14880 |
| | tatttttagtt | taaaagttaa | taaatatttt | gttagaaatt | ccaatctgct | tgtaatttat | 14940 |
| | caataaacia | aatattaaat | aacaagctaa | agtaacaaat | aatatcaaac | taatagaaac | 15000 |
| | agtaatctaa | tgtaacaaaa | cataatctaa | tgctaataata | acaaagcgca | agatctatca | 15060 |

| | | | | | | | |
|----|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| | ttttatatag | tattatttttc | aatcaacatt | cttattaatt | tctaaataat | acttgtagtt | 15120 |
| | ttattaactt | ctaaatggat | tgactattaa | ttaaataaat | tagtcgaaca | tgaataaaca | 15180 |
| | aggtaacatg | atagatcatg | tcattgtgtt | atcattgatc | ttacatttgg | attgattaca | 15240 |
| | gttgggaagc | tgggttcgaa | atcgataaagc | ttggatcctc | tagaccactt | tgtacaagaa | 15300 |
| 5 | agctgaacga | gaaacgtaaa | atgatataaa | tatcaatata | ttaaattaga | ttttgcataa | 15360 |
| | aaaacagact | acataatact | gtaaaacaca | acatatccag | tcactatgaa | tcaactactt | 15420 |
| | agatgggtatt | agtgacctgt | agtcgactaa | gttggcagca | tcacccgacg | cactttgcgc | 15480 |
| | cgaataaata | cctgtgacgg | aagatcactt | cgcagaataa | ataaatcctg | gtgtccctgt | 15540 |
| | tgataccggg | aagccctggg | ccaacttttg | gcgaaaatga | gacgttgatc | ggatttcaca | 15600 |
| 10 | actcttatac | ttttctctta | caagtcgttc | ggcttcatct | ggatttttcag | cctctatact | 15660 |
| | tactaaacgt | gataaagttt | ctgtaatttc | tactgtatcg | acctgcagac | tggctgtgta | 15720 |
| | taagggagcc | tgacatttat | attccccaga | acatcaggtt | aatggcggtt | ttgatgtcat | 15780 |
| | tttcgcggtg | gctgagatca | gccacttctt | ccccgataac | ggagaccggc | acactggcca | 15840 |
| | tatcgggtggt | catcatgcgc | cagctttcat | ccccgatatg | caccaccggg | taaagttcac | 15900 |
| 15 | gggagacttt | atctgacagc | agacgtgcac | tggccagggg | gatcaccatc | cgctgcccgg | 15960 |
| | gcgtgtcaat | aatatcactc | tgtacatcca | caaacagacg | ataacggctc | tctcttttat | 16020 |
| | aggtgtaaac | cttaaactgc | atttcaccag | tccctgttct | cgtcagcaaa | agagccgttc | 16080 |
| | atttcaataa | accggggcgc | ctcagccatc | ccttccctgat | tttccgcttt | ccagcggttcg | 16140 |
| | gcacgcagac | gacgggcttc | attctgcatg | gttgtgttta | ccagaccgga | gatattgaca | 16200 |
| 20 | tcatatatgc | cttgagcaac | tgatagctgt | cgctgtcaac | tgtcactgta | atacgtgtgt | 16260 |
| | tcatagcaca | cctctttttg | acatacttct | gttcttgatg | cagatgattt | tcaggactat | 16320 |
| | gacactagcg | tatatgaata | ggtagatgtt | tttattttgt | cacacaaaaa | agaggctcgc | 16380 |
| | acctcttttt | cttatttctt | tttatgattt | aatacggcat | tgaggacaat | agcgagtagg | 16440 |
| | ctggatacga | cgattccgtt | tgagaagaac | atttggaagg | ctgtcgggtc | actaagttgg | 16500 |
| 25 | cagcatcacc | cgaagaacat | ttggaaggct | gtcggtcgac | tacaggtcac | taataccatc | 16560 |
| | taagtagttg | attcatagtg | actggatag | ttgtgtttta | cagtattatg | tagtctgttt | 16620 |
| | tttatgcaaa | atctaattta | atataattgat | atttataatca | ttttacgttt | ctcgttcagc | 16680 |
| | ttttttgtac | aaacttgtct | agagtcctgc | tttaatgaga | tatgcgagac | gcctatgatc | 16740 |
| | gcatgatatt | tgctttcaat | tctgttgtgc | acgttgtaaa | aaacctgagc | atgtgtagct | 16800 |
| 30 | cagatcctta | ccgccgggtt | cgggttcattc | taatgaatat | atcaccgggt | actatcgtat | 16860 |
| | ttttatgaat | aatattctcc | gttcaattta | ctgattgtac | cctactactt | atatgtacaa | 16920 |
| | tattaaaatg | aaaacaatat | attgtgctga | ataggtttat | agcgacatct | atgatagagc | 16980 |
| | gccacaataa | caaacaattg | cgttttatta | ttacaaatcc | aatttttaaa | aaagcggcag | 17040 |
| | aaccgggtcaa | acctaaaaga | ctgattacat | aaatcttatt | caaatttcaa | aaggccccag | 17100 |
| 35 | gggctagtat | ctacgacaca | ccgagcggcg | aactaataac | gttactgaa | gggaactccg | 17160 |
| | gttccccgcc | ggcgcgcatg | ggtgagattc | cttgaagttg | agtattggcc | gtccgctcta | 17220 |
| | ccgaaagtta | cgggcaccat | tcaaccgggt | ccagcacggc | ggccgggtta | ccgacttgct | 17280 |
| | gccccgagaa | ttatgcagca | tttttttggt | gtatgtgggc | cccaaataaa | gtgcaggtca | 17340 |
| | aaccttgaca | gtgacgacaa | atcggtgggc | gggtccaggg | cgaattttgc | gacaacatgt | 17400 |
| 40 | cgaggctcag | caggacctgc | aggcatgcaa | gctagcttac | tagtgatgca | tattctatag | 17460 |
| | tgtcacctaa | atctgc | | | | | 17476 |

We claim:

- 1) A method for monitoring the reduction of the expression of a target gene in a cell of a eukaryotic organism, said method comprising the steps of
 - a) providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region
 - i) said first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of said target gene;
 - ii) said second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from said sense nucleotide region of the target gene ;
 - iii) said first region and said second region being capable of forming a double stranded RNA region;
 - iv) said third region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene present in said eukaryotic cell and which is different from said target gene;
 - v) said fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of said 19 consecutive nucleotides from said sense nucleotide region of said second gene;
 - vi) said third and fourth region being capable of forming a double stranded RNA region; and
 - b) monitoring the reduction of the expression of said target gene by analyzing the reduction in expression of said second gene.
- 2) The method of claim 1, wherein said eukaryotic organism is a plant.

- 3) The method of claim 1, wherein said eukaryotic organism is an animal.
- 4) The method of claim 1, wherein said eukaryotic organism is a yeast, fungus
5 or mold.
- 5) The method of claim 1 or claim 2, wherein eukaryotic organism is selected
from the group of cotton, potato, corn, wheat, rice, sugar cane, oilseed rape,
Arabidopsis, sugarbeet, tobacco and soybean.
10
- 6) The method of claim 1 or claim 3, wherein said eukaryotic organism is
selected from the group consisting of insects, shellfish, molluscs,
crustaceans, crabs, lobsters, prawns, fish, avian animals, mammals and
humans.
15
- 7) The method of any one of claims 1 to 6, wherein said second gene is an
endogenous gene present in said eukaryotic cell.
- 8) The method of any one of claims 1 to 6, wherein said second gene is a
20 transgene stably integrated into the genome of said eukaryotic cell.
- 9) The method of claim 2 or claim 5, wherein said second gene is selected
from the group consisting of PDS, EIN2, FLC and PhyB.
- 25 10) The method of any one of claims 1 to 6, wherein said eukaryotic cell
comprises a functionally expressed GUS or a GFP gene, and said second
gene is a GUS or GFP gene.
- 11) The method of any one of claims 1 to 10, wherein said first and second
30 region, and said third and fourth region are about 300 nt in length.

- 12) The method of any one of claims 1 to 11, wherein said dsRNA is transcribed from a chimeric gene comprised within cells of said eukaryotic organism, said chimeric gene comprising:
- a) a promoter region which functions in said eukaryotic cell; operably linked to
 - b) a DNA region which when transcribed yields said dsRNA molecule; and
 - c) a transcription termination and polyadenylation region that functions in said eukaryotic organism.
- 13) The method according to claim 12, wherein said chimeric gene is stably integrated into the genome of cells of said eukaryotic organism.
- 14) An RNA molecule as described in any one of claims 1 to 13.
- 15) Use of an RNA molecule according to claim 14 for measuring the reduction of expression of a target gene.
- 16) A DNA molecule for measuring the reduction of expression of a target gene in a cell of a eukaryotic organism comprising
- a) a promoter region which functions in said eukaryotic cell; operably linked to
 - b) a DNA region which when transcribed yields a dsRNA molecule, said dsRNA comprising a first, second, third and fourth region;
 - i) said first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of said target gene;
 - ii) said second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from said sense nucleotide region of the target gene ;

- iii) said first region and said second region being capable of forming a double stranded RNA region:
- iv) said third region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene present in said eukaryotic cell and which is different from said target gene;
- 5 v) said fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of said 19 consecutive nucleotides from said sense nucleotide region of said second gene;
- 10 vi) said third and fourth region being capable of forming a double stranded RNA region; and
- c) a transcription termination and polyadenylation region that functions in said eukaryotic organism, wherein said second gene is an endogenous gene of said eukaryotic organism or a transgene stably integrated into the genome of cells of said eukaryotic organism.
- 15
- 17) Use of a DNA molecule according to claim 16 to measure the expression of said target gene by measuring the reduction in expression of said second gene.
- 20
- 18) A non-human eukaryotic organism comprising an RNA molecule according to claim 14
- 25
- 19) A non-human eukaryotic organism comprising a DNA molecule according to claim 16.
- 20) The non-human eukaryotic organism according to claim 18 or 19, which is a plant.
- 30

21) The non-human eukaryotic organism according to claim 18 or 19, which is an animal.

22) The non-human eukaryotic organism according to claim 18 or 19, which is a yeast, fungus or mold.

23) A method for identifying, within a population of dsRNA-mediated gene-silenced eukaryotic organisms, the organisms with the desired degree of silencing of a target gene comprising:

10

a) providing cells of said eukaryotic organisms with a dsRNA comprising a first region, a second region, a third region and a fourth region

15

i) said first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of said target gene;

20

ii) said second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to a nucleotide sequence complementary to 19 consecutive nucleotides from said sense nucleotide region of the target gene ;

25

iii) said first region and said second region being capable of forming a double stranded RNA region:

iv) said third region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from a sense nucleotide region of a second gene present in said eukaryotic cell and which is different from said target gene;

30

v) said fourth region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of said 19 consecutive nucleotides from said sense nucleotide region of said second gene;

- vi) said third and fourth region being capable of forming a double stranded RNA region; and
- b) identifying said organism with said desired degree of silencing of said target gene, by selecting said organisms with the desired degree of silencing of said second gene.

- 24) A method for modulating the reduction of the expression of a target gene in a cell of a eukaryotic organism, comprising the steps of
- a) providing the eukaryotic cell with a dsRNA comprising a first region, a second region, a third region and a fourth region,
 - i) said first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of said target gene;
 - ii) said second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from said sense nucleotide sequence of said target gene;
 - iii) said first region and said second region being capable of forming a double stranded RNA region;
 - iv) said third region and said fourth region comprising complementary nucleotide sequences which have a sequence identity of less than 50% to the nucleotide sequence of said target gene and which are capable of forming a double stranded RNA,
- wherein said target gene is an endogenous gene in said eukaryotic cell or a transgene stably integrated in the genome of said eukaryotic cell.

- 25) The method of claim 24, wherein the size of the double stranded RNA capable of being formed by base-pairing between said third and fourth region being equal or larger than the size of the double stranded RNA

capable of being formed by base-pairing between said first and said second region.

26)The method of claim 24 or 25, wherein said eukaryotic organism is a plant.

5

27)The method of claim 24 or 25, wherein said eukaryotic organism is an animal.

28)The method of claim 24 or 25, wherein said eukaryotic organism is a yeast,
10 fungus or mold.

29)The method of claim 26, wherein said plant is selected from the group of cotton, potato, corn, wheat, rice, sugar cane, oilseed rape, *Arabidopsis*, sugarbeet, tobacco or soybean.

15

30)The method of claim 27, wherein said animal is selected from the group consisting of insects, shellfish, molluscs, crustaceans, crabs, lobsters, prawns, fish, avian animals, mammals and humans.

20 31)An RNA molecule for modulating the expression of a target gene in a cell of a eukaryotic organism, comprising a first region, a second region, a third region and a fourth region,

i) said first region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to 19 consecutive nucleotides from the sense nucleotide sequence of said target gene;

25

ii) said second region comprising a nucleotide sequence of at least 19 consecutive nucleotides having at least 94% sequence identity to the complement of 19 consecutive nucleotides from said sense nucleotide sequence of said target gene;

30

- iii) said first region and said second region being capable of forming a double stranded RNA region;
- iv) said third region and said fourth region comprising complementary nucleotide sequences which have a sequence identity of less than 50% to the nucleotide sequence of said target gene and which are capable of forming a double stranded RNA,
- wherein said target gene is an endogenous gene in said eukaryotic cell or a transgene, stably integrated in the genome of said eukaryotic cell.
- 32) A DNA molecule, capable of producing a dsRNA molecule according to claim 31 comprising a DNA region which when transcribed yields said dsRNA molecule, said DNA region being operably linked to a promoter and a transcription termination and polyadenylation signal.
- 33) Use of an RNA molecule according to claim 31, to modulate the reduction of expression of a target gene in a eukaryotic organism.
- 34) Use of a DNA molecule according to claim 32, to modulate the reduction of expression of a target gene in a eukaryotic organism.
- 35) A non-human eukaryotic organism comprising a DNA molecule according to claim 32.
- 36) A non-human eukaryotic organism comprising an RNA molecule according to claim 31.

1/2

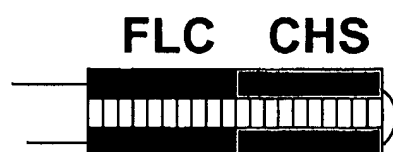
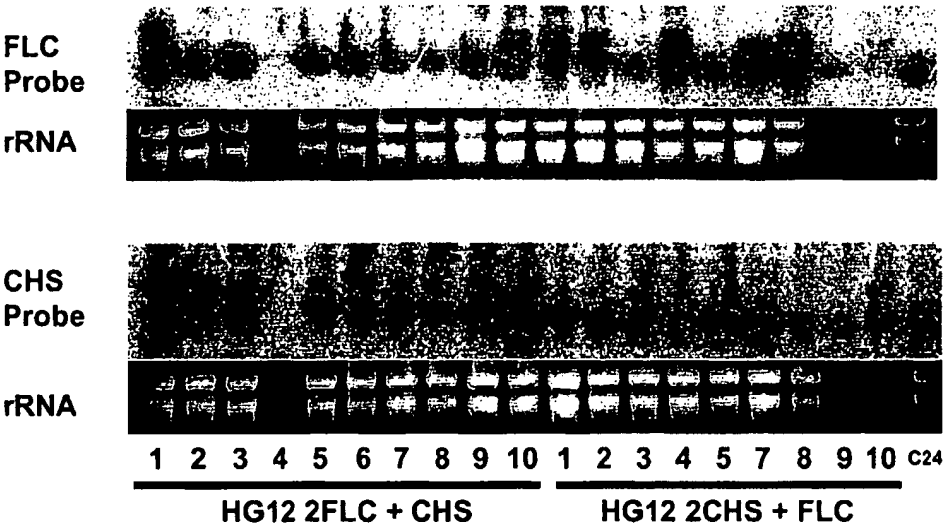
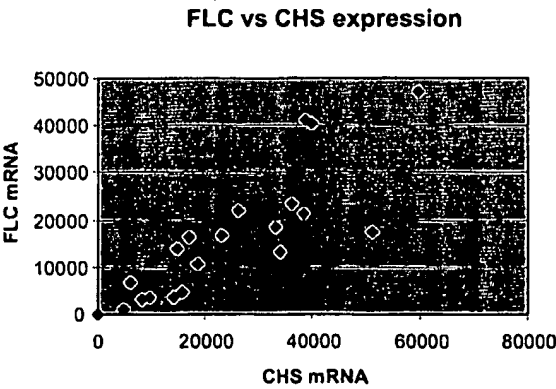


Figure 1



A.



B.

Figure 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU03/00293

| | | |
|---|--|--|
| A. CLASSIFICATION OF SUBJECT MATTER | | |
| Int. Cl. ⁷ : C12N 15/11, C12Q 1/68 | | |
| According to International Patent Classification (IPC) or to both national classification and IPC | | |
| B. FIELDS SEARCHED | | |
| Minimum documentation searched (classification system followed by classification symbols) SEE ELECTRONIC DATABASE BOX BELOW | | |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched SEE ELECTRONIC DATABASE BOX BELOW | | |
| Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) WPIDS, Medline, CA, Biosis: RNAi, RNA interfer, RNA or gene silencing, reporter, 19, 21, dicer | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT | | |
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| A | ELBASHIR, S. M. et al (2002 Feb) "Analysis of gene function in somatic mammalian cells using small interfering RNAs" <i>Methods</i> , 26:199-213. See entire document | |
| <input type="checkbox"/> Further documents are listed in the continuation of Box C <input type="checkbox"/> See patent family annex | | |
| <p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p> | | |
| Date of the actual completion of the international search 30 April 2003 | | Date of mailing of the international search report 5 - MAY 2003 |
| Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929 | | Authorized officer PHILIPPA WYRDEMAN Telephone No : (02) 6283 2554 |